Université Paris Descartes

Ecole doctorale Gc2iD

Spécialité : Génétique

Thèse pour l'obtention du grade de Docteur de l'Université Paris Descartes

Présentée et soutenue par:

Magalie VATIN

Le 01 octobre 2012

Étude de gènes impliqués dans la fertilité humaine à partir d'un modèle de souris interspécifiques recombinantes congéniques (IRCS)

Jury :

Pr Marc DELPECH	Président
Dr Olivier SANDRA	Rapporteur
Dr Gérard CHAOUAT	Rapporteur
Dr Corinne COTINOT	Examinateur
Dr Florian GUILLOU	Examinateur
Dr Ahmed ZIYYAT	Directeur de thèse
Dr Catherine SERRES	Invitée
Dr Daniel VAIMAN	Invité

UNIVERSITÉ PARIS DESCARTES
FACULTÉ DE MÉDECINE

Université Paris Descartes

Ecole doctorale Gc2iD

Spécialité : Génétique

Thèse pour l'obtention du grade de Docteur de l'Université Paris Descartes

Présentée et soutenue par:

Magalie VATIN

Le 01 octobre 2012

Etude de gènes impliqués dans la fertilité humaine à partir d'un modèle de souris interspécifiques recombinantes congéniques (IRCS)

Thèse dirigée par le Dr Ahmed ZIYYAT

Département Génétique et Développement, Equipe D. VAIMAN
« Génomique, Epigénétique et Physiopathologie de la Reproduction »
Institut Cochin, Inserm U1016,
24 , rue du Faubourg St Jacques, 75014 Paris

Pour toi maman,

Et comme dirait l'autre :

"Sa ki la pouw, dlo pa ka chayé'y"

"Bondyé pli fò ki dyab"

Remerciements

Je voudrais tout d'abord remercier le Dr Olivier SANDRA et le Dr Gérard CHAOUAT d'avoir accepté d'être les rapporteurs de ma thèse. Merci également aux Drs Corinne COTINOT et Florian GUILLOU, examinateurs et au Pr Marc DELPECH, président du jury.

En 2008, j'ai intégré l'équipe «Génomique, Epigénétique et Physiopathologie de la Reproduction» à l'institut Cochin, à l'occasion de la réalisation de mon stage de Master 2. Je tiens donc à remercier tout particulièrement Daniel VAIMAN d'avoir accepté de m'encadrer, avec David L'Hôte, au cours de mon stage de Master, d'avoir eu confiance en moi et de m'avoir présentée au concours de l'école doctorale Gc2ID, ce qui m'a permis d'effectuer ma thèse au sein de son équipe. Je remercie Catherine SERRES et Ahmed ZIYYAT, mes deux directeurs de thèse. En effet, j'ai débuté sous la direction de Catherine qui m'a dirigée pendant deux ans puis a transmis cette responsabilité à Ahmed en raison de son départ en retraite. Ces deux chercheurs ont su me transmettre leur savoir, m'écouter, me guider et m'accompagner dans mon travail de recherche. Avec Ahmed notre collaboration a été aussi le début d'une grande complicité et d'une réelle amitié. Il me parait important de souligner que même en étant à la retraite Catherine a été très présente jusqu'à la fin de ma thèse et je lui en suis très reconnaissante.

Au sein de l'équipe de Daniel VAIMAN, j'ai eu la chance de découvrir des personnalités marquantes au contact desquelles j'ai gagné en maturité. Je tiens donc à remercier tous les chercheurs en particulier Paul LAISSUE, Reiner VEITIA et le Pr Marc FELLOUS pour leurs précieux conseils, les ingénieurs, les techniciens, les thésards et les

étudiants que j'ai eu la chance de côtoyer durant ces quatre années. Un clin d'œil à Jana AUER et Brigitte LEFEVRE qui m'ont écoutée et soutenue quand j'en avais besoin, Carmen MARCHIOL et Gilles RENAULT de PIPA pour leur soutien, Isabelle LANCTIN et Jennifer CHEVALIER de l'Institut Pasteur pour les souris (sans souris pas de thèse !!!), Mariem OUNI et Sylvie BOUVIER que j'ai eu le plaisir d'encadrer à l'occasion de la réalisation de leur stage de Master 2.

Je tiens à remercier tous les professeurs qui ont cru en moi et qui m'ont encouragée, en particulier le Dr Laurence GALL de l'INRA de Jouy en Josas qui m'a transmis la passion de la recherche au cours de mon stage de Master 1.

Je remercie toute ma famille, je vous aime, maman mon inspiration et ma force, ma sœur Valérie l'exceptionnelle, Cindy notre sœur de cœur et tous mes amis en particulier ma petite Ines.

Pour finir, cette thèse je la dédie au courage de ma mère et à la mémoire de mon père parti trop tôt...

Table des matières

Liste des abréviations

AMH	Anti-müllerian hormone
AMP	Assistance Médicale à la Procréation
RSA	Avortements spontanés à répétition
AZF	Azoospermia factor
B6	C57BL6/J
CSS	Chromosome Substitution Strains
FCS	Fausse couche spontanée
FIV	Fécondation In Vitro
FSH	Hormone Folliculo-stimulante
IA	Insémination artificielle
ICSI	Injection intra-cytoplasmique de spermatozoïde
IOP	Insuffisance ovarienne prématurée
IRCS	Lignées interspécifiques recombinantes congéniques
KO	Knock-out
LH	Hormone lutéinisante
MFR	Monthly Fecundity Rates
OMS	Organisation Mondiale de la Santé
QTL	Quantitative Trait Locus
RCS	Recombinant Congenic Strains
RFLP	Restriction Fragments Length Polymorphysm
RIS	Recombinant Inbreed Strains

SA	Semaines d'Aménorrhée
SAPL	Syndrome Anti-PhosphoLipide
SDP	Strain Distribution Pattern
SEG	SEG/P
SNP	Single Nucleotide Polymorphism
SRY	Sex determining Region of the Y chromosome
TDF	Facteur de détermination testiculaire
TTP	Time-To-Pregnancy

Introduction

I. La fertilité des mammifères

La reproduction sexuée est un processus physiologique majeur nécessaire à la survie de toutes les espèces de mammifères. Elle nécessite l'interaction d'un spermatozoïde (gamète mâle) et d'un ovocyte (gamète femelle) pour former un zygote ce qui constitue la première étape dans le développement de l'embryon. La fabrication de ces deux gamètes est intimement liée à la reproduction sexuée.

Le déterminisme du sexe est a priori déterminé génétiquement dès la fécondation, lors de la mise en commun du patrimoine génétique des gamètes mâle et femelle et, à l'exception de quelques espèces chez les mammifères, il est basé sur l'existence des chromosomes sexuels X et Y. C'est au niveau du bras court du chromosome Y qu'est localisé le facteur de détermination testiculaire (TDF) codé par le gène *SRY* (Sex Determining Region of the Y chromosome) cloné en 1990 dans le laboratoire de Peter Goodfellow.

La différenciation de l'ébauche gonadique (ou crête génitale) et le développement de l'appareil génital et des organes sexuels en testicules ou en ovaires, constituant ainsi le sexe gonadique, dépend de la présence ou de l'absence de SRY. (Jacobs and Strong, 1959; Russell, 1961; Sekido and Lovell-Badge, 2008). En réponse à des signaux hormonaux, le tractus génital et les

organes génitaux externes appropriés se développent ensuite pour définir le sexe phénotypique. Ce sont essentiellement les hormones testiculaires (testostérone et AMH ou anti-müllerian hormone) qui sont à l'origine du choix de la différenciation du sexe apparent (Ravel et al., 2004).

Au cours de la différenciation du testicule, SRY et SOX9 régulent négativement la voie de différenciation ovarienne par la répression de RSPO1. Parallèlement, dans la gonade femelle, en l'absence de SRY, le haut niveau de RSPO1 entraine l'activation de la voie femelle.

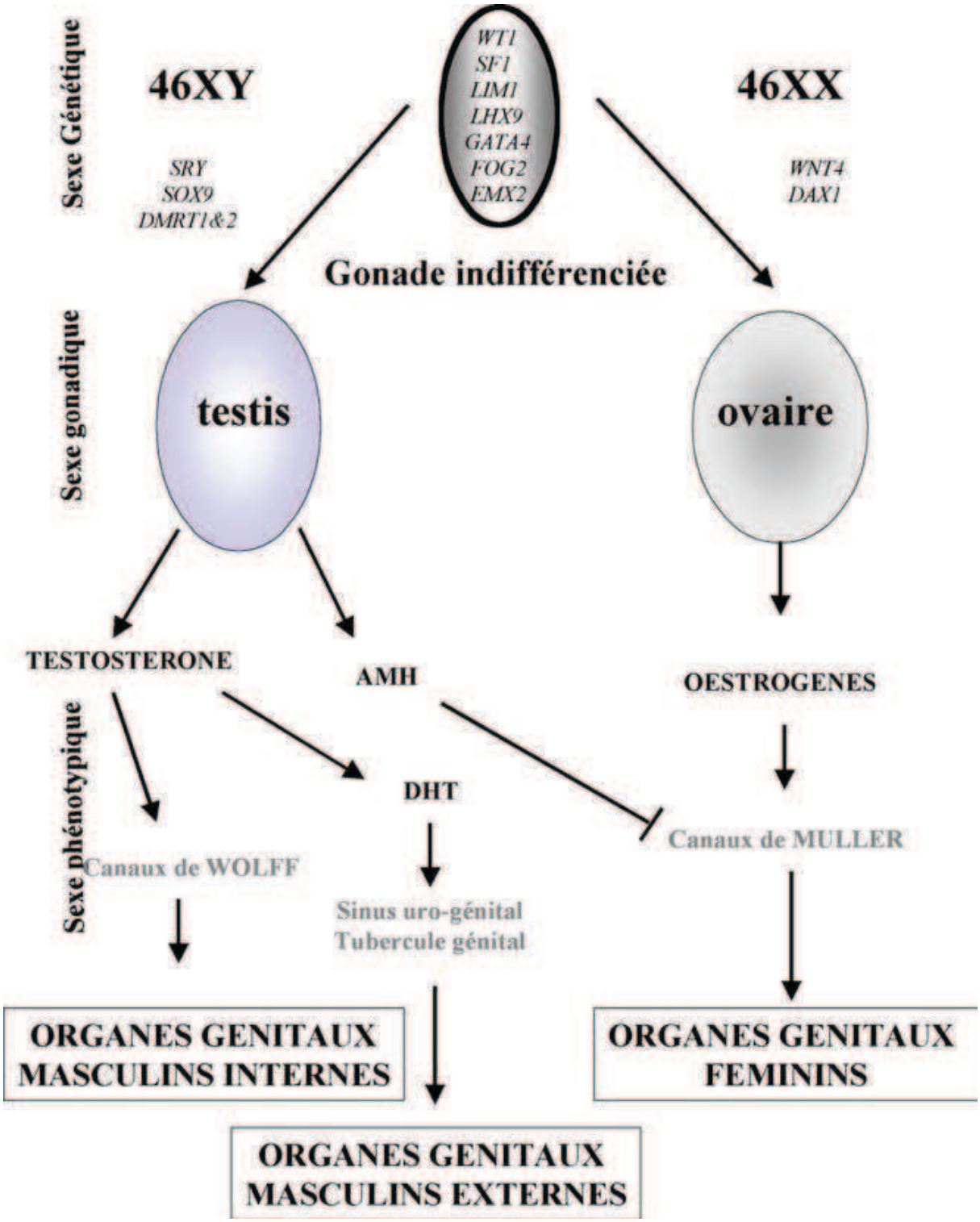


***Schéma récapitulatif de la cascade d'activation génique aboutissant au déterminisme gonadique puis à la différenciation sexuelle (d'après* C. Ravel et al. 2004).**

De nombreux travaux essentiellement réalisés chez la souris ont permis de mettre en évidence les évènements morphogénétiques impliqués dans la formation d'un ovaire ou d'un testicule.

Les étapes se déroulant dans l'ovaire (folliculogenèse, ovulation), la fécondation, l'implantation et le développement embryonnaire précoce sont finement régulés génétiquement (Figure 1). De même, le grand nombre de gènes spécifiquement exprimés dans la lignée germinale mâle illustre la complexité du processus de la spermatogenèse et indique que des mutations dans des milliers de gènes différents pourraient provoquer l'infertilité masculine (Tableau 1) (Matzuk and Lamb, 2008; Schultz et al., 2003). Il est fort probable que cette complexité contribue à un grand nombre d'infertilité idiopathique chez l'Homme.

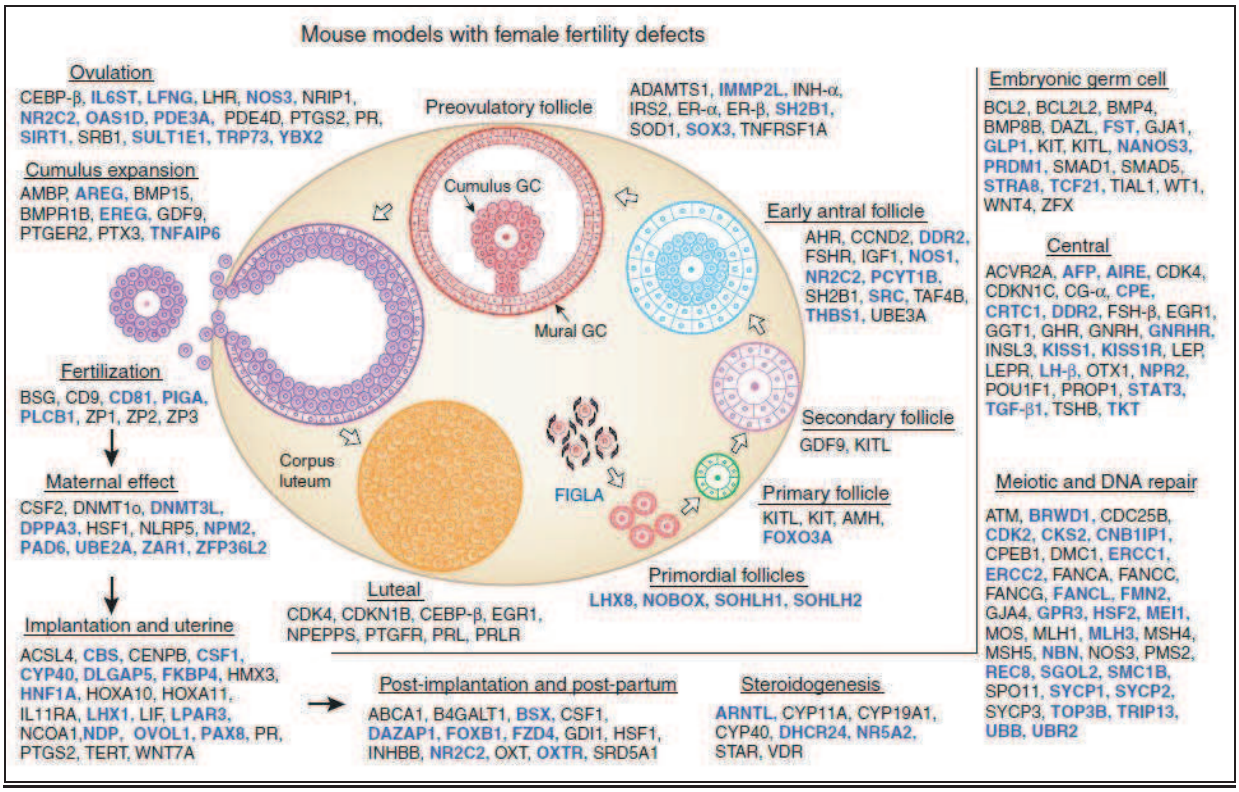


Figure 1 : Les protéines impliquées dans la fertilité femelle identifiées chez la souris. (Matzuk and Lamb, 2008). *Des modèles de souris KO (knock-out) ont permis d'identifier des protéines clés impliquées dans la formation de follicules, la folliculogenèse ovarienne, l'ovulation, et les événements post-ovulatoires (la fécondation, l'implantation et le développement embryonnaire précoce). L'expression de ces gènes est finement régulée et sous le contrôle du système hypothalamo-hypophysaire. L'étude de plus d'une centaine de mutants dans les espèces murines et ovines a donc permis d'accroître les connaissances concernant la reproduction femelle. Cependant, ces données ne permettent pas d'expliquer les causes génétiques d'infertilité féminine qui restent largement inexpliquées. Les gènes présentés en bleu représentent l'accroissement des connaissances sur le sujet (et donc du nombre de gènes connus comme étant impliqués) entre une revue de 2002 et celle de 2008.*

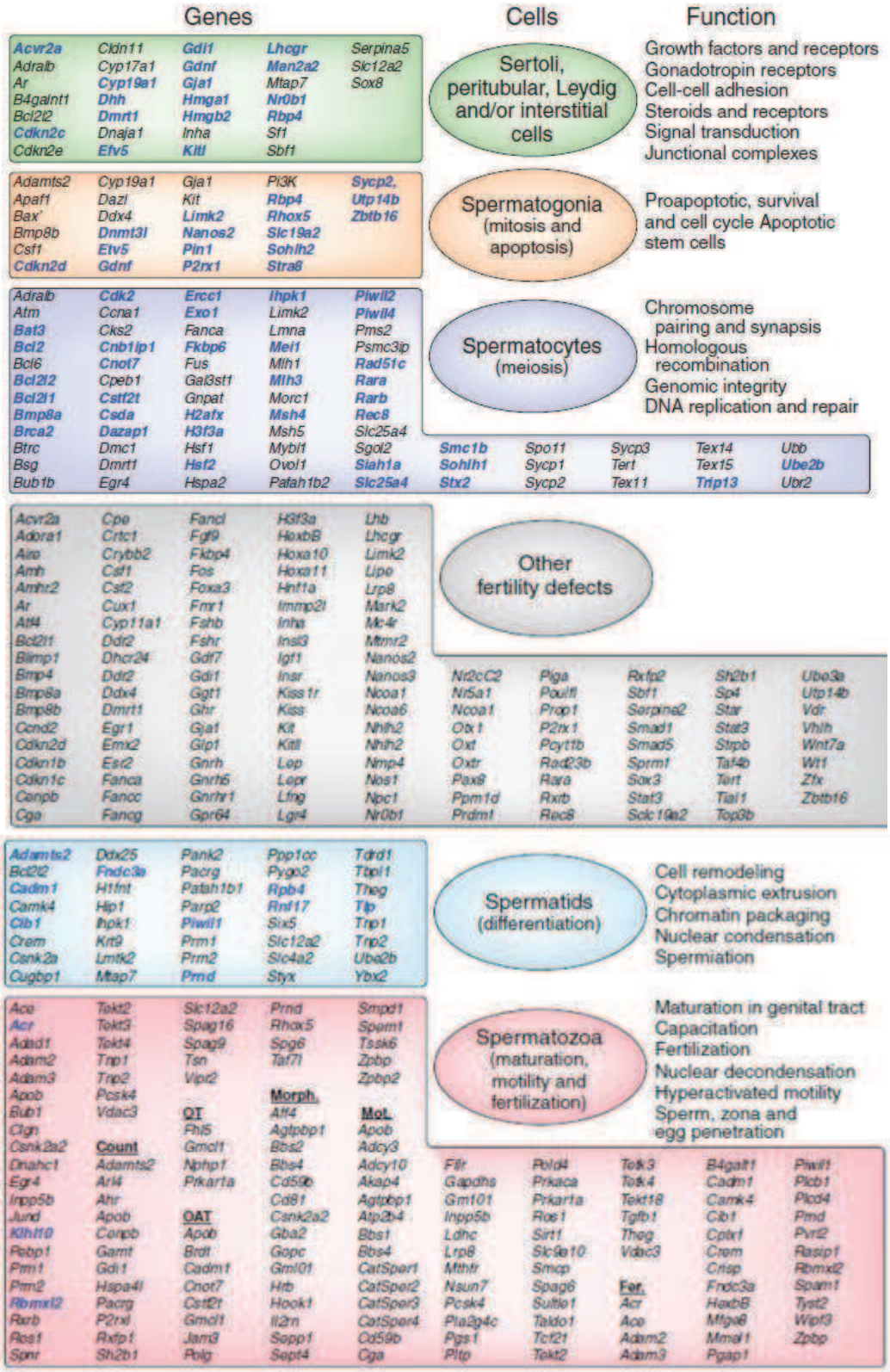


Table 1: Gènes impliqués dans l'infertilité masculine mis en évidence par le modèle murin (Matzuk and Lamb, 2008) *Les gènes présentés en bleu représentent l'accroissement des connaissances sur le sujet (et donc du nombre de gènes connus comme étant impliqués) entre une revue de 2002 et celle de 2008.*

II. L'infertilité humaine

A. Définition et données épidémiologiques

La reproduction est une fonction physiologique qui comprend de nombreuses étapes, telles que la formation des gamètes, la fécondation, le développement préimplantatoire, l'implantation et le développement post-implantatoire. Ainsi l'aptitude à procréer dépend du bon fonctionnement coordonné des systèmes reproducteurs mâle et femelle.

Les épidémiologistes et les démographes ont coutume de distinguer fertilité/infertilité, de fécondité/infécondité. La fertilité correspond à une aptitude, celle d'un couple à concevoir; elle exprime une probabilité. Au contraire, la fécondité est un état, celui d'un couple qui a conçu; elle exprime un constat (d'après la revue du Pr Alfred Spira, « Infertilité – infécondité », 1997). Par définition, la fertilité est la capacité de produire une descendance, alors que la fécondité est la capacité biologique à se reproduire sur la base de la probabilité mensuelle de conception. Un couple est dit infertile s'il ne parvient pas à mener une grossesse à terme après un an ou plus de rapports sexuels réguliers non protégés (Gnoth et al., 2005; Zegers-Hochschild et al., 2006b).

La fécondité peut être mesurée par le temps pris pour obtenir une grossesse. On parle de «Time-To-Pregnancy» (TTP) exprimée en «Monthly Fecundity Rates» (MFR), c'est à dire la probabilité d'obtenir une grossesse dans un cycle menstruel. La MFR chez l'homme est relativement faible (20%) par rapport aux autres espèces de mammifères (Evers, 2002; Stevens, 1997; Teklenburg et al., 2010; Viudes-de-Castro and Vicente, 1997). On parle d'infertilités modérées et sévères définies par des MFR de 5 et 1%, respectivement (Evers, 2002). À l'autre extrémité du spectre, la superfertilité peut être caractérisée par une MFR de 60% ou plus. Sur la base du modèle de Tietze (Evers, 2002; Tietze et al., 1950), 79% de la population est estimée fertile, 18% infertile et 3% superfertile.

L'OMS (Organisation Mondiale de la Santé) considère comme infertile un couple qui, au bout de deux ans de relations sexuelles, sans contraception, ni intervention médicale, n'a pas conçu. On parle d'infertilité quand au moins un des membres du couple est stérile ou quand sa capacité à concevoir est sérieusement diminuée. On distingue l'infertilité primaire quand il n'y a jamais eu de grossesse et l'infertilité secondaire s'il y a eu une ou plusieurs grossesses à terme. Dans l'espèce humaine, la fécondité ou probabilité mensuelle d'obtenir une grossesse est d'environ 25%. Dans le cas d'une incapacité totale à procréer naturellement on parle de stérilité. Pour une probabilité de procréer <5% on parle d'hypofertilité (OMS). L'infertilité est un problème de santé publique

majeur qui concerne entre 8 et 12% des couples à l'échelle mondiale soit plus de 80 millions de personnes, avec, d'après l'OMS, 40 % de couples atteints d'infertilité primaire et 60 % d'infertilité secondaire (Evers, 2002).

Les causes d'infertilité ont très longtemps été attribuées aux femmes. Ce n'est que depuis le début du 20ème siècle que la piste masculine est explorée. Dans environ 30% des cas, il est trouvé une cause purement féminine, dans 20% des cas une cause uniquement masculine, dans 40% des cas une cause qui incombe aux deux partenaires et dans 10% des cas l'infertilité reste inexpliquée.

Ainsi, en France 60 000 couples (~1/6) consultent chaque année pour une difficulté à concevoir. Ces couples sont pris en charge par des services de santé de la reproduction et ils se voient proposer des traitements et/ou une assistance médicale à la procréation (AMP). En 2003, une naissance française sur 20 a été obtenue à l'issue d'un traitement ou d'une technique médicale. Cependant le taux de succès des techniques d'AMP reste faible, de l'ordre de 20 % par tentative (Andersen and Loft, 2006). Une des techniques d'AMP les plus connues est la Fécondation In Vitro (FIV) mais d'autres techniques existent (Insémination Artificielle (IA), Injection Intra-Cytoplasmique de Spermatozoïde (ICSI). Le premier enfant issu de la technique de FIV date de 1978 (Steptoe and Edwards, 1978). De plus la proportion de naissances obtenues par cette technique n'a cessé de progresser au cours des vingt dernières années, passant

de 0,52 % des enfants conçus par FIV en 1988 à 1,74 % en 2006 (soit un rythme moyen de hausse de + 0,72 % par décennie) (de La Rochebrochard et al., 2008).

En dehors de toute pathologie, des facteurs liés au mode de vie, ont un rôle établi dans la modulation de la fertilité. Chez la femme, l'âge joue un rôle très important puisque son augmentation diminue les chances de grossesse et dans le même temps augmente le risque de fausses couches. De la même façon, l'obésité, le tabac et l'alcool diminuent la fertilité des femmes (Caserta et al., 2011; Kelly-Weeder and Cox, 2006; Norman et al., 2004). Par exemple, les études de Hughes en 1996 et Augood en 1998 ont démontré que le tabagisme chez les femmes diminue de façon significative les chances de conception. Chez l'homme, l'âge est également associé à une baisse de la fertilité mais à un degré bien moindre par rapport à la femme. L'obésité, le tabac et l'alcool sont également des facteurs délétères à la fertilité masculine. Certaines conditions de l'environnement (chaleur, solvants, pesticides...) peuvent aussi entraîner une diminution de la fécondité (Homan et al., 2007), de même que des facteurs de risque professionnels, y compris les travaux physiques pénibles et l'exposition aux gaz anesthésiques, les métaux lourds et les solvants (Homan et al., 2007; Kumar, 2004). Beaucoup moins d'attention a été portée à l'exposition aux produits chimiques présents dans l'environnement et l'alimentation (Foster et al., 2008), en dépit de preuves provenant d'études toxicologiques qui indiquent

clairement que de nombreux produits chimiques nuisent à la reproduction des femelles à la fois directement et/ou indirectement (Caserta et al., 2011).

B. Les causes masculines d'infertilité

1. Généralités

Chez l'homme, le spermogramme associé au spermocytogramme constituent les examens de base dans le cadre de l'exploration de la fertilité.

Les normes établies par l'OMS sont les suivantes : volume entre 2 et 6 ml, pH entre 7,4 et 8, concentration en spermatozoïdes entre 20 et 200 millions/ml, vitalité supérieure à 60 %, mobilité supérieure à 50% avec au moins deux tiers de spermatozoïdes mobiles progressifs («fléchants»), et formes morphologiquement normales supérieures à 40 % avec un index d'anomalies multiples inférieur à 1,6.

L'infertilité masculine est un syndrome multifactoriel englobant une grande variété de troubles. Cependant, pour plus de la moitié des hommes infertiles, la cause est inconnue et peut être congénitale ou acquise (Bhasin et al., 1997; Poongothai et al., 2009). L'enquête Thonneau, a permis de répertorier les causes principales d'infertilité masculine (Thonneau et al., 1991).

L'oligo-térato-asthénozoospermie représente la cause principale d'infertilité masculine (21%).

L'oligospermie (2%) correspond à une diminution du nombre de spermatozoïdes (moins de 20 millions/ml). L'azoospermie (9%) correspond à une absence totale de spermatozoïde. Il peut s'agir soit d'une absence de production par les testicules (azoospermie sécrétoire (6%)) soit d'une obturation des canaux excréteurs (azoospermie obstrusive (3%)).

L'asthénospermie (ou asthénozoospermie) (17%) correspond à un défaut de mobilité (moins de 40%) ou de mouvement des spermatozoïdes (vitesse, trajectoire). Ces anomalies peuvent être dues à des anomalies de structure des spermatozoïdes. La nécrospermie correspond à un pourcentage élevé de spermatozoïdes morts (>50%) le plus souvent dû à des infections. La tératospermie (ou tératozoospermie) (10%) est la présence d'un taux anormalement élevé de spermatozoïdes anormaux (supérieur à 85%). Le pourcentage minimal de spermatozoïdes normaux dans un sperme normal varie entre 15 et 50%. Ces anomalies peuvent intéresser toutes les parties du spermatozoïde (tête, flagelle) et sont généralement dues à un dysfonctionnement de la spermiogenèse, la dernière phase de la spermatogenèse.

Concernant les causes génétiques de l'infertilité masculine, des études dans les années 1970, notamment celle de Tiepolo L et Zuffardi O en 1976, ont

conduit à suspecter l'existence d'un facteur putatif AZF *(Azoospermia factor)* codé par un ou plusieurs gènes, indispensable(s) à la spermatogenèse et situé(s) en région Yq11, ultérieurement subdivisée en 3 sous-régions non chevauchantes dénommées (AZFa, AZFb et AZFc) (Tiepolo and Zuffardi, 1976). Environ 10 à 20% des hommes, dépourvus de spermatozoïdes dans leurs éjaculats présentent une délétion au niveau du locus AZF (Azoospermia Factor) qui serait responsable d'atteintes plus ou moins sévères de la spermatogenèse. Plusieurs gènes candidats ont été trouvés à l'intérieur de ces régions. Dans les cas d'azoospermie, les gènes de la famille DAZ, situés dans la région AZFc, sont les candidats les plus souvent délétés (Navarro-Costa et al., 2010). Récemment, une nouvelle région du chromosome Y de 1,6 Mb nommée «gr/gr » a été décrite comme délétée en particulier chez les hommes infertiles avec des degrés divers de défauts spermatiques. Cependant, la compréhension des causes génétiques de l'infertilité masculine reste limitée (Poongothai et al., 2009).

2. La spermatogenèse, une étape clé de la fertilité mâle

La majorité des anomalies de fertilité masculine sont liés à des défauts de la spermatogenèse, un phénomène long et complexe qui débute à la puberté et se poursuit pendant toute la vie d'un individu de sexe masculin.

La spermatogenèse est l'événement physiologique majeur qui se produit dans les tubes séminifères, les unités fonctionnelles des testicules, et qui permet

de produire des spermatozoïdes. Chaque homme produit plus de 100 millions de spermatozoïdes chaque jour sous l'influence de l'hormone folliculo-stimulante (FSH) et de l'hormone lutéinisante (LH) (Sharpe et al., 1994). La FSH exerce ses effets sur les cellules de Sertoli dans l'épithélium séminifère pour assurer le maintien de la spermatogenèse (Griswold et al., 1995). La LH se lie à ses récepteurs au niveau des cellules de Leydig pour induire la stéroïdogenèse et permettre la production de la testostérone et de l'œstradiol (Carreau and Hess, 2010; O'Donnell et al., 2001). Ces deux hormones hypophysaires (FSH et LH) qui régulent la spermatogenèse sont sous l'influence de l'hypothalamus. La relation hormonale et fonctionnelle complexe, entre l'hypothalamus, l'hypophyse et le testicule, est connue sous le nom de l'axe hypothalamo-hypophyso-testiculaire. Cet axe est notamment la cible de la contraception hormonale masculine (Anderson and Baird, 2002; Liu et al., 2010; Nieschlag, 2010; Nieschlag et al., 2011).

En 2006, Misell et al. ont confirmé qu'un cycle spermatogénétique, de l'initiation de la spermatogenèse à l'apparition des spermatozoïdes dans l'éjaculat, dure environ 64 jours (Misell et al., 2006). La spermatogenèse se décompose en trois grandes phases:

1- Pendant la première phase, les spermatogonies (ou cellules souches germinales) prolifèrent par mitoses (capacité d'auto-renouvellement) et se différencient en spermatocytes I qui s'engagent dans le processus de la méiose,

2- La seconde phase correspond donc à la méiose, processus de division des cellules germinales diploïdes, permettant d'obtenir des gamètes haploïdes, les spermatides essentielles pour la reproduction sexuée.

3- La troisième phase est la spermiogénèse qui correspond à la transformation des spermatides rondes en spermatides allongées puis en spermatozoïdes dotés de la capacité de motilité (après phagocytose du corps résiduel par les cellules de Sertoli). La spermiogénèse permet la formation de l'acrosome et induit des changements nucléaires, le développement du flagelle, la réorganisation du cytoplasme et des organites cellulaires (Figure 2).

Le processus de libération du spermatozoïde par la cellule de Sertoli (à partir de l'épithélium dans la lumière tubulaire) est appelé spermiation (Figure 3).

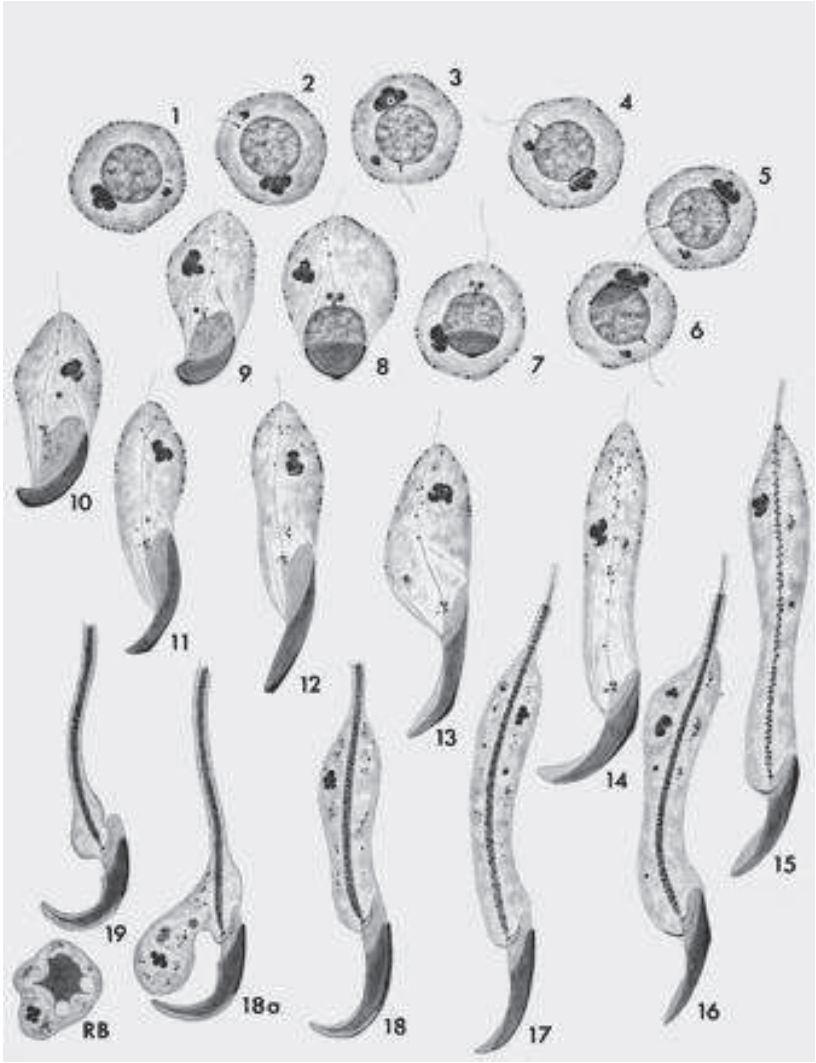


Figure 2 : Illustration de la spermiogenèse du rat adulte applicable à d'autres mammifères (Clermont and Rambourg, 1978). *Les numéros des dessins correspondent aux étapes de la spermiogénèse telles que proposées par Leblond en 1952 (Leblond and Clermont, 1952). Au cours des trois premières étapes de la spermiogenèse, l'appareil de Golgi juxtanucléaire élabore les granules proacrosomiques, qui fusionnent en un seul granule acrosomique qui s'attache au noyau. Durant les étapes 4 à 7, l'appareil de Golgi contribue à la formation de l'acrosome qui se positionne à la surface du noyau. Pendant cette phase les mitochondries sont situées le long de la membrane plasmique, et les centrioles liés au noyau au pôle opposé de l'acrosome. Cela va donner lieu à l'axonème du flagelle en croissance. Durant les étapes 8 à 14, la spermatide devient polarisée, la chromatine se condense, et le noyau s'allonge pour prendre sa forme caractéristique du modèle murin, en faucille. Au cours des étapes de 15 à 19, phase de maturation, le flagelle termine sa formation par la migration des mitochondries vers le segment proximal en développement pour former la gaine mitochondriale. Enfin, le lobe cytoplasmique apical et la majeure partie du cytoplasme se séparent de la spermatide qui devient, après élimination du corps résiduel (RB), un spermatozoïde.*

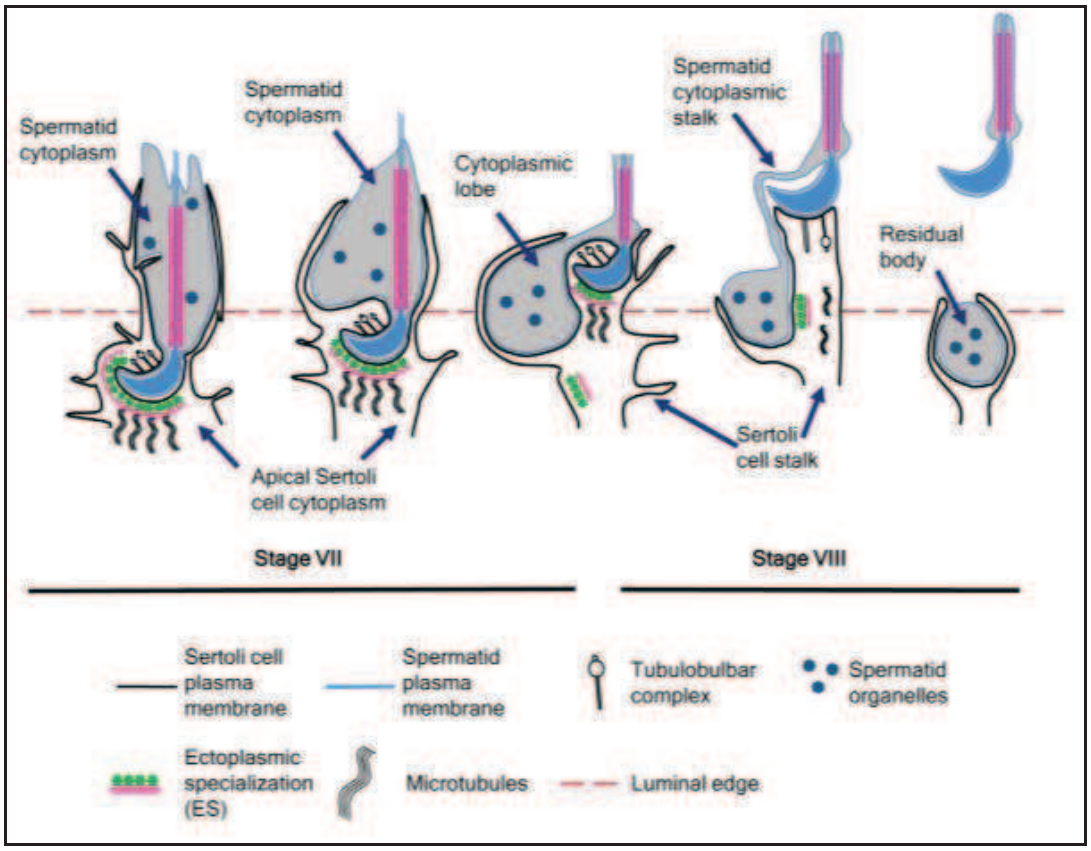


Figure 3 : Le processus de spermiation (O'Donnell et al., 2011). *Avant la spermiation, la spermatide allongée interagit avec les cellules de Sertoli par l'intermédiaire d'une vaste structure connue sous le nom de la spécialisation ectoplasmique (ES). Au cours de la spermiation, les spermatides allongées montrent des changements progressifs dans la morphologie nucléaire et cytoplasmique. Au début de la spermiation au stade VII, la spermatide montre un cytoplasme abondant autour du flagelle, et est en grande partie enveloppé par des projections du cytoplasme apical des cellules de Sertoli (parfois appelé le processus apical). Les complexes tubulo-bulbaires se forment dans la courbure ventrale de la tête des spermatides, dans les zones déficitaires en ES. La tête et le flagelle de la spermatide sont progressivement attirés plus loin dans la lumière tubulaire alors que son cytoplasme se condense et reste immobile dans l'épithélium, ce qui a pour effet de former le lobe cytoplasmique où se concentrent les organites cytoplasmiques, qui finira par former le corps résiduel. Au cours de la progression du stade VII au stade VIII, la structure ES disparaît de la membrane plasmique de la cellule de Sertoli opposée à la tête de la spermatide. Les complexes tubulo-bulbaires montrent également des changements dans leur morphologie au cours de la progression du processus de spermiation, ils sont plus nombreux dans la courbure ventrale de la tête de la spermatide au stade VII. La spermiation se termine par la libération de la tête de la spermatide des cellules de Sertoli et la séparation de la spermatide et du corps résiduel qui sera phagocyté par les cellules de Sertoli.*

Au terme de la spermatogenèse, le gamète possède toutes ses structures morphologiques indispensables à la fécondation. Toutefois il est important de noter que le spermatozoïde testiculaire subit, aussi bien dans les voies génitales masculines que dans les voies génitales féminines, de nombreuses modifications physiologiques nécessaires à l'accomplissement des étapes précoces de la fécondation, à savoir le franchissement des cellules du cumulus, la reconnaissance et la pénétration de la zone pellucide puis l'adhésion et la fusion des membranes plasmiques des deux gamètes. La complexité des événements cytologiques à la base de la formation des spermatozoïdes supposent une multitude de mécanismes de régulation, biochimiques et physico-chimiques (Cheng et al., 2011).

L'identification des gènes jouant un rôle dans la spermatogenèse est principalement basée sur l'étude des modèles murins, en particulier le modèle de la souris. Chez l'homme, une telle identification reste largement inaccessible en raison de la difficulté à avoir accès à des échantillons de tissus à partir d'une gonade mâle et les possibilités limitées de vérification fonctionnelle. Par conséquent, les causes génétiques d'infertilité masculine sont encore mal connues. Cependant l''identification des gènes potentiellement impliqués dans la spermatogenèse doit être améliorée afin d'accroître l'efficacité des tests de diagnostic indispensables à la caractérisation des causes de stérilité chez les

hommes afin de leurs proposer un traitement médical ciblé ou une assistance médicale adéquate (Waclawska and Kurpisz, 2012).

C. Les causes féminines d'infertilité

1. Généralités

L'enquête Thonneau, en 1991, a permis de répertorier les causes principales d'infertilité féminine (Thonneau et al., 1991).

Chez la femme, les troubles de l'ovulation constituent la cause principale d'infertilité (33 à 35%). Si l'ovulation est complètement absente on parle d'anovulation ou si elle est présente mais de mauvaise qualité on parle de dysovulation. Certaines anomalies des ovaires sont à l'origine de troubles ovulatoires particuliers tels que le Syndrome des ovaires micro- polykystiques (présence de nombreux petits kystes). D'autres dysfonctionnements peuvent être la conséquence d'un épuisement prématuré des ovaires. C'est le cas de l'insuffisance ovarienne prématurée (IOP ou ménopause précoce) correspondant à une défaillance ovarienne survenant avant l'âge de 40 ans.

Dans 26% des cas, des anomalies des trompes (obturation), pour la plupart dues à des infections génitales, peuvent être un obstacle mécanique à la rencontre des gamètes dans les voies génitales. Elles peuvent aussi causer des grossesses extra-utérines.

De même, les malformations utérines (bicornes unicervicaux 39%, cloisonnés 34%, bicornes bicervicaux 11%, unicornes 5%) concernent 3,5% des femmes infertiles.

L'endométriose est une pathologie concernant 4% des cas d'infertilité féminine. Elle correspond à la prolifération anormale de la muqueuse utérine en dehors de l'utérus, en particulier au niveau des ovaires pouvant perturber la rencontre de l'ovocyte et des spermatozoïdes et influer sur la qualité de l'ovulation.

Enfin la glaire cervicale peut également être en cause (dans 4% des cas d'infertilité féminine), elle peut être secrétée de façon inadéquate (quantité, qualité, pH) à cause d'infections, de dysfonctionnement des glandes endocervicales ou suite à des traitements de lésions du col de l'utérus.

2. Les fausses couches spontanées répétées

a) Données épidémiologiques:

On parle de fausse couche spontanée (FCS) en cas « d'expulsion ou d'extraction hors de la mère d'un embryon ou d'un fœtus pesant moins de 500g et de moins de 22 semaines d'aménorrhée (SA) » (OMS, 1997). On parle de fausse couche « précoce » lorsque celle-ci survient avant 15 SA. Douze à 15%

des grossesses cliniquement reconnues et 50 à 60% de toutes les grossesses se terminent par un AS dans l'espèce humaine (Edmonds et al., 1982; Wilcox et al., 1988) pour la majeure partie très tôt dans la gestation, autour de la période d'implantation. Il y a une augmentation du risque d'avortement en fonction du nombre d'accidents précédents (Figure 4 et Tableau 2) (Knudsen et al., 1991; Regan et al., 1989)

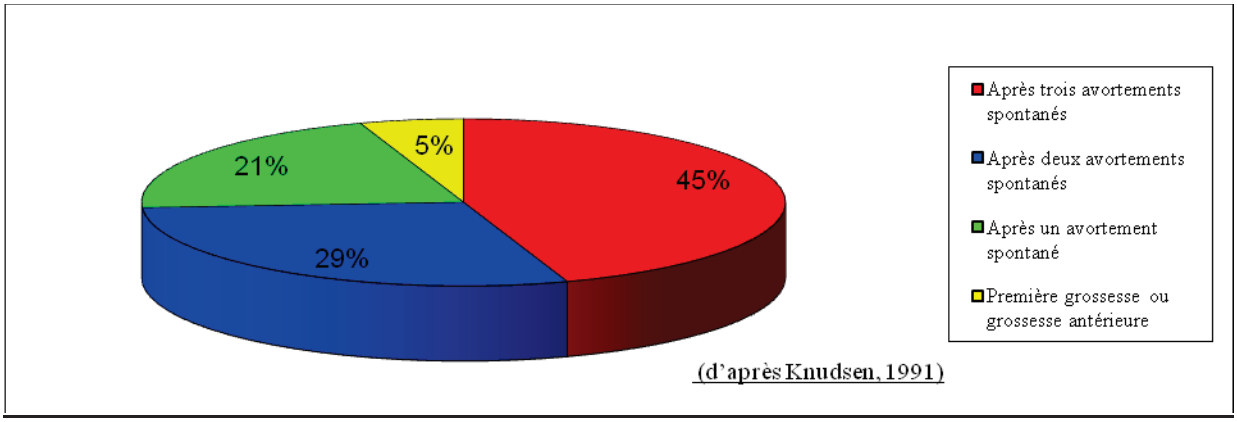


Figure 4 : Augmentation du risque d'AS avec le nombre d'antécédents (Knudsen et al., 1991) Le risque d'un AS est déterminé par l'issue des grossesses antérieures.

Nombre de fausses couches	Risque de FCS pour la grossesse en
0	10-15%
1	13-26%
2	17-35%
3	25-46%

Tableau 2 : Evaluation du risque de récidive en fonction du nombre de fausses couches spontanées antérieures (« Protocoles cliniques en obstétrique » Par Dominique Cabrol, François Goffinet, paru le 13/11/2008)

On définit les avortements spontanés à répétition (RSA) comme au moins trois avortements successifs. Un à 5% des couples seraient concernés par l'RSA

(Rai and Regan, 2006). Cependant les causes des RSA restent inconnues dans 50 à 80% des cas (2008; Plouffe et al., 1992).

Bien que la qualité de l'implantation puisse être à l'origine d'RSA pendant toute la durée de la gestation, diverses raisons font penser que la pathogénèse est bien plus complexe. Le risque d'un nouvel avortement est directement lié au succès des gestations précédentes (Knudsen et al., 1991; Regan et al., 1989).

b) Etiologie des RSA

Les causes des RSA peuvent être anatomiques avec une fréquence des malformations congénitales variant entre 2 et 38%. Le cloisonnement utérin est la malformation utérine la plus souvent associée aux RSA (Salim et al., 2003).

Un avortement précoce est dans la plupart des cas associé à un caryotype anormal de l'embryon. Les anomalies chromosomiques les plus souvent en cause sont les trisomies autosomiques, la monosomie X et les polyploïdies. Trois pour cent des couples confrontés à des AS comportent un conjoint présentant des réarrangements chromosomiques importants. L'anomalie chromosomique la plus fréquemment rencontrée est la translocation balancée (transfert de segments entre chromosomes non homologues) (Stephenson and Sierra, 2006).

Les causes immunologiques peuvent être auto-immunes (syndrome anti-phospholipide ou SAPL) ou allo-immunes (réponse immunitaire inappropriée envers l'embryon pouvant diminuer la tolérance maternelle). La prévalence du SAPL chez des femmes atteintes d'RSA est d'environ 15% (Rai et al., 1995). Dans la plupart des études on relève une augmentation des anticorps anti-phospholipides allant jusqu'à 42% chez les femmes stériles. Les anticorps anti-phospholipides pourraient intervenir à la fois sur la nidation et sur la croissance fœtale. Les pertes fœtales seraient la conséquence d'une ischémie placentaire (Balasch et al., 1998; Roussev et al., 1998). Les traitements du SAPL combinent l'aspirine et l'héparine, et améliorent la survie des fœtus. En effet, 70% des patientes traitées ont des gestations qui sont menées à terme (Empson et al., 2002). Il a également été montré que l'hyperactivation du système du complément due à une formation exagérée de complexes immunologiques, est fréquente chez des patientes sujettes au SAPL (Oku et al., 2012).

Des causes endocriniennes peuvent également expliquer certains cas d'RSA. Le diabète mal soigné, le syndrome des ovaires polykystiques et l'hyper-prolactinémie sont associés à un risque d'RSA. L'administration de bromocriptine, un inhibiteur de la sécrétion de prolactine, diminue le taux d'avortements (Hirahara et al., 1998). De plus la prévalence des ovaires polykystiques est de 40% chez des patientes atteintes d'RSA tandis que chez des

femmes sans antécédent elle n'est que de 22% (Rai et al., 2000). D'autres études montrent également que la résistance à l'insuline est associée à une augmentation des RSA (Craig et al., 2002).

III. Introduction à la génétique quantitative

A. Les caractères quantitatifs complexes

La reproduction, comme de nombreux processus physiologiques (ou pathologiques), fait intervenir des paramètres quantitatifs et complexes résultant de l'interaction entre des facteurs génétiques et des facteurs environnementaux. L'infertilité humaine se manifeste en effet dans de nombreux cas par différents degrés d'hypofertilité plutôt que par un état d'infertilité absolue.

Un caractère quantitatif est un caractère phénotypique dont la distribution est une fonction continue et mesurable. Un caractère qualitatif peut être monogénique, c'est-à-dire lié à la variation allélique d'un seul gène, ou polygénique/oligogénique, c'est-à-dire contrôlé par l'association de différents allèles d'un nombre limité de gènes. La génétique quantitative est aussi appelée génétique multifactorielle, ou génétique des caractères à déterminisme complexe. Dans ces cas, différents locus agissent sur le phénotype, chacun de façon modérée (et non en tout ou rien) en interaction entre eux (de la simple additivité des effets, à des relations plus complexes), et dont l'action peut être modulée par l'environnement. Dans le cas de l'additivité, le génotype confère au phénotype une valeur moyenne égale à la somme de l'effet de chaque variant

allélique. A l'opposé, on parle d'épistasie lorsque l'interaction des génotypes multiplie les effets avec une valeur phénotypique moyenne différente de la somme des effets des variants alléliques (Carlborg and Haley, 2004).

Pour l'étude des caractères quantitatifs au sein d'une population, on définit la variabilité dite intra-individuelle comme étant liée à l'expérimentateur et à l'environnement et la variabilité interindividuelle qui persistera entre des individus génétiquement identiques dans un environnement contrôlé. Une mesure quantitative se traduit par une moyenne et un écart type calculés au sein d'une population.

L'association statistique d'un caractère quantitatif avec une région du génome traduit l'existence d'un QTL (Quantitative Trait Locus). Un QTL se définit comme un locus polymorphe pour lequel au moins deux formes alléliques sont associées à des valeurs phénotypiques moyennes différentes pour le caractère étudié. Un caractère quantitatif peut également être sous le contrôle d'un ou plusieurs QTL(s) qui contribuent de façon plus ou moins importante à l'élaboration du phénotype. On parle de « force » du QTL définie comme étant le pourcentage de variance du phénotype qu'il explique. La force du QTL sera déterminante pour sa détection, qui sera d'autant plus aisée que celle-ci sera grande.

B. Principe de détection des QTL

Lors de l'élaboration d'un protocole expérimental visant à la détection de QTL, le choix des lignées parentales est déterminant. A priori, la probabilité de détecter des QTL sera d'autant plus importante que le polymorphisme phénotypique sera marqué entre les deux lignées. Le succès de la cartographie initiale des QTL est donc principalement basé sur l'exploitation de lignées parentales fortement polymorphes.

Le modèle murin est l'outil principal pour l'étude de caractères quantitatifs. Les premiers QTL ont été cartographiés chez la souris dans les années 90 (Hilbert et al., 1991). La cartographie de QTL se fait en deux étapes. La première étape est la primo-localisation des QTL correspondant au screening global du génome. La seconde étape est la cartographie fine permettant de proposer des gènes candidats expliquant la variation phénotypique.

La détection de QTL nécessite la connaissance précise de la carte génétique de l'espèce étudiée correspondant au positionnement de marqueurs génétiques polymorphes les uns par rapport aux autres le long d'un chromosome. Au cours de l'histoire du clonage de QTL, en fonction des avancées techniques, trois types de marqueurs moléculaires ont été utilisés pour l'essentiel : les RFLP (Restriction Fragments Length Polymorphysm), les microsatellites et les SNP (Single Nucleotide Polymorphism).

La technique RFLP est basée sur l'existence d'un polymorphisme de longueur de fragments de restriction. En effet, le nombre de sites de coupure pour des enzymes de restriction et leurs positions dans le génome diffèrent en fonction des individus.

Les microsatellites sont de très courtes séquences d'ADN (1 à 4 nucléotides) répétées en tandem susceptibles de varier en nombre de copies en fonction du fond génétique.

Les SNP (Single Nucleotide Polymorphism) sont des variations d'une seule paire de base du génome entre individus d'une même espèce, à une position déterminée. En moyenne, on en compte un tous les 500 bases d'ADN chez l'Homme ; comme les autres marqueurs présentés ici ils se rencontrent dans les séquences codantes et non codantes.

La détection de QTL peut être basée sur une analyse par marqueur qui consiste à comparer pour un marqueur donné les contrastes phénotypiques correspondant aux deux allèles du marqueur, et à évaluer leur significativité. Si la comparaison des moyennes révèle une différence statistiquement significative, on en déduit la présence d'un QTL au voisinage de ce marqueur (Lander and Botstein, 1989). L'analyse point par point permet de calculer un LOD score qui mesure la vraisemblance de l'existence d'un QTL pour tous les marqueurs qui ont été analysés le long d'un chromosome.

LOD score = log (P(H1)/P(H0)) avec P(H1) = Probabilité d'obtenir la déviation phénotypique sous l'hypothèse H1 (existence du QTL à proximité du marqueur considéré) et P(H0)= Probabilité sous l'hypothèse H0 (absence de QTL à proximité de ce marqueur)

Du fait des spécificités génétiques de l'espèce murine (lignées consanguines en particulier), la souris a été au centre de l'identification de QTL chez les mammifères. Trois grands types de lignées de laboratoire ont été particulièrement utilisés pour la cartographie de QTL chez la souris :

(1) Les CSS (Chromosome Substitution Strains) sont des lignées pour lesquelles un chromosome entier est remplacé par son homologue d'une autre lignée. Ces lignées permettent une cartographie au chromosome prés.

(2) Les RIS (Recombinant Inbreed Strains) sont obtenues à partir de deux lignées éloignées phylogénétiquement pour créer une population F2 qui va subir une vingtaine de générations de croisements frère-sœur afin de fixer chaque locus à l'état homozygote. Les lignées RIS possèdent en moyenne 50% du génome des deux lignées parentales. Le phénotypage d'un ensemble de lignée RIS (chacune correspondant à une combinaison unique des génomes parentaux) suivi de la comparaison de leur pattern de distribution d'haplotypes permet de déterminer le fragment génomique minimal responsable d'une variation phénotypique, c'est l'analyse SDP (Strain Distribution Pattern).

(3) Les RCS (Recombinant Congenic Strains) sont obtenues par l'établissement d'une population F1, croisée en retour sur l'une des lignées parentales (receveuse). La génération N2 obtenue est croisée sur le même parent. La mise à l'état consanguin se fait à partir des individus N3 pendant au moins 20 générations. En moyenne, le pourcentage de génome donneur dans chaque lignée est théoriquement de 12,5%. De ce fait, la probabilité de mettre en évidence des relations d'épistasie entre deux QTL est deux fois plus importante avec les RCS comparée au RIS (Figure 5).

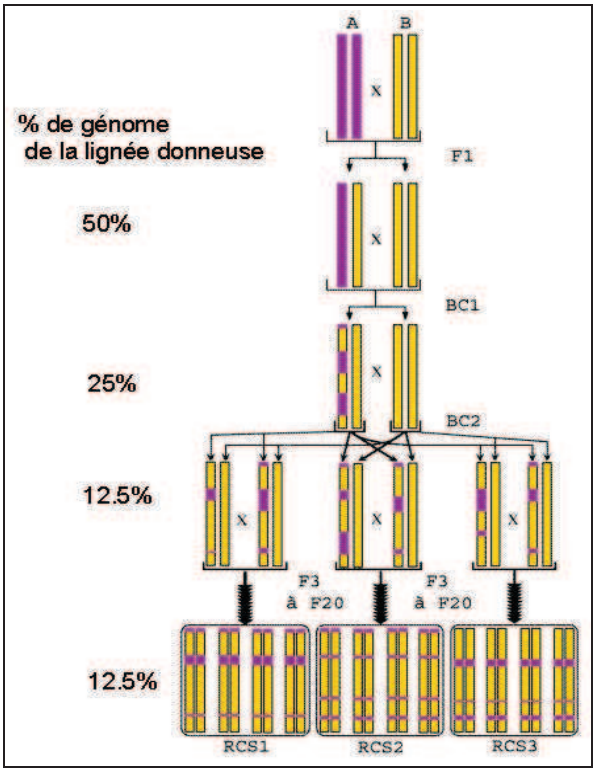


Figure 5 : Etablissement du set de RCS (Demant and Hart, 1986). *Ces lignées sont obtenues à l'issue de 2 backcross (ou croisement en retour) et possède 12,5% de génome donneur et 77,5% de génome receveur. 20 lignées RCS permettent une couverture théorique du génome de 93%.*

Dans ce travail de thèse nous nous intéresserons à un modèle RCS particulier, les lignées interspécifiques recombinantes congéniques (IRCS). La particularité de ces lignées réside dans le fait qu'elles ont été créées à partir d'un croisement dans lequel le génome donneur et le génome receveur appartiennent à deux espèces de souris différentes, ce qui confère aux croisements un fort polymorphisme.

C. Les IRCS pour la cartographie de QTL

Les lignées interspécifiques recombinantes congéniques (IRCS) ont été créées à l'institut Pasteur à partir d'un croisement initial entre deux lignées parentales C57BL6/J (B6) et SEG/P (SEG) appartenant à deux espèces différentes, *Mus musculus domesticus* et *Mus spretus* respectivement (Burgio et al., 2007). Ce croisement interspécifique entre souris de laboratoire et souris sauvages proposé par Jean-Louis Guenet et Xavier Montagutelli a abouti à l'obtention de lignées consanguines possédant un polymorphisme phénotypique et génotypique très important. En effet ces deux espèces ont divergé il y a

environ 1,5 à 2 millions d'années et présentent donc un polymorphisme de séquence proche de celui qui existe entre l'homme et le chimpanzé (Guenet and Bonhomme, 2003; Newman et al., 2005).

Plus précisément, le premier croisement entre les lignées parentales a été suivi de 2 croisements en retour (backcross) sur B6 donnant naissance à la génération N2. La génération N3 a été obtenue à l'issue d'une vingtaine de croisements consanguins frère-sœur. Le résultat est la création d'un set de 55 lignées congéniques, contenant en moyenne 2% de génome SEG fixé à l'état homozygote dans un fond génétique B6 (Figure 6). De façon intéressante, il faut noter que la proportion théorique (12.5%) est largement supérieure à celle réellement obtenue. La tolérance du génome d'une espèce à celui d'une autre espèce pourrait être limitée en quantité, voire en qualité : on ne peut exclure que certaines régions du génome *spretus* ne soient totalement incompatibles avec d'autres régions du génome *musculus*. Il s'ensuit que l'approche IRCS ne peut probablement pas permettre de détecter des QTL sur tout le génome.

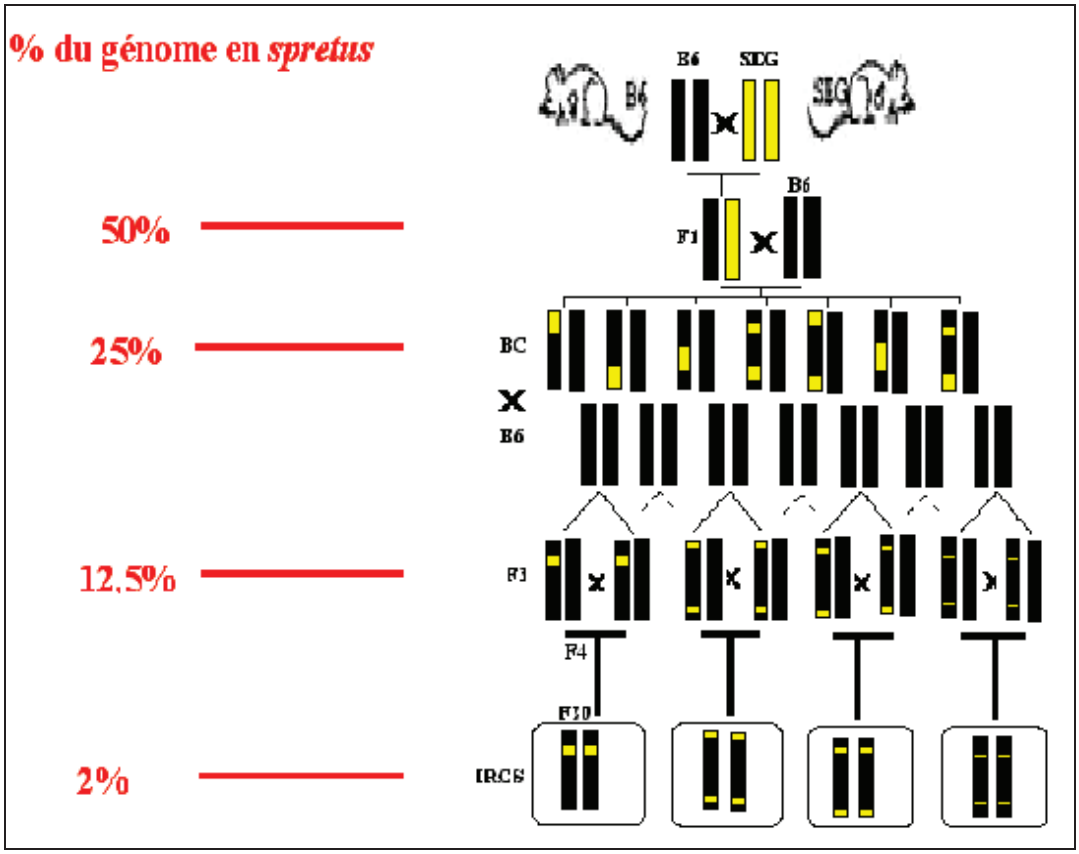


Figure 6 : Etablissement des lignées IRCS créées à l'institut Pasteur. *Le premier croisement entre le lignées parentales B6 et SEG est suivi par deux backcross sur B6. Une vingtaine de croisements frère – sœur permettent de fixer le génome spretus à l'état homozygote dans un fond B6. Chacune des 55 lignées IRCS possède en moyenne 1,5% du génome spretus.*

Les lignées IRCS ainsi obtenues ont été génotypées par microsatellites et SNP, en collaboration avec le centre national de génotypage d'Evry ce qui a permis une couverture totale du génome des lignées. Les fragments *spretus* (11,7 Mb en moyenne) répartis entre les différentes lignées permettent une couverture du génome d'environ 37,8% avec en moyenne 3,15 fragment par lignée. Cependant il est important de noter que les chromosomes X et Y ne sont pas couverts. La faible couverture du génome est liée à la perte de plus de la

moitié des lignées (63 lignées perdues sur 118) lors de la réalisation des croisements consanguins, perte principalement dues à une forte contre-sélection des fragments *spretus* (Figure 7).

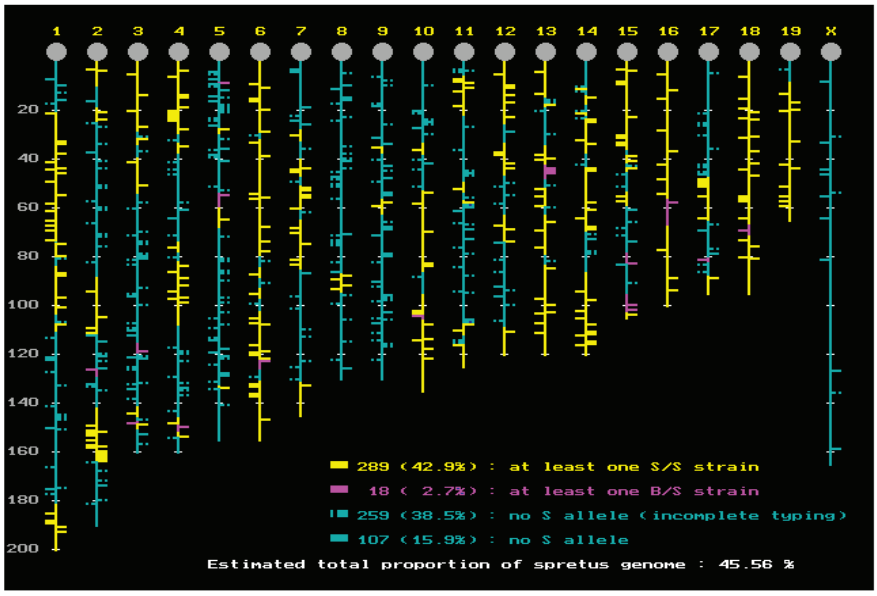


Figure 7: Carte génétique composite de l'ensemble du set d'IRCS créées à l'Institut Pasteur-Paris (http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm). *En jaune sont représentées les régions homozygotes spretus présentent dans au moins une lignée IRCS et en en rose, à l'état hétérozygote. En bleu, sont représentées les régions musculus.*

Le fort polymorphisme entre les lignées SEG et B6 permet la cartographie fine de QTL impliqués dans la variation de caractères quantitatifs. C'est ainsi qu'au cours de ce travail de thèse et au cours des années précédentes, l'exploitation de ces lignées a permis d'identifier des QTL, des phénomènes d'épistasie entre QTLs mais également de proposer des gènes candidats responsables de défauts de fertilité mâle ou femelle.

IV. Apport du modèle IRCS pour la cartographie de QTL de Fertilité

Depuis maintenant près d'une dizaine d'années le modèle IRCS est exploité au sein de notre équipe dans le but d'identifier de nouveaux QTL de fertilité. Deux études qui constituent le point de départ de mon travail de thèse, concernent respectivement des phénotypes de fertilité mâle et femelle.

A. Cartographie de QTL de la fertilité mâle

La première étude a été publiée en 2007. Ces travaux avaient pour but la primo-localisation de locus régulant la fertilité mâle par le phénotypage du set de 53 lignées IRCS et des deux lignées parentales sur la base de paramètres dont la mesure fut simple et rapide (L'Hote et al., 2007).

- Poids des organes reproducteurs (testicules, épididymes, vésicules séminales et prostate)
- Production des gamètes (mesure de la réserve spermatique épididymaire totale)

– Qualité des spermatozoïdes (morphologie, mouvement, réaction acrosomique)

Les mesures sur les lignées IRCS ont été comparées à celle de la lignée de référence B6. Des lignées présentant des déviations phénotypiques significatives ont été identifiées pour la quasi-totalité des phénotypes mesurés. Parmi les résultats notables, le phénotypage a permis notamment de mettre en évidence une réduction du poids testiculaire dans la lignée 97C (moins 30% par rapport à B6) (Figure 8) et une fréquence relativement élevée de tératozoospermie (environ 20%, Figure 9).

Cette lignée 97C qui présente la plus forte réduction du poids testiculaire observée dans les IRCS, possède deux fragments *spretus*, le premier sur le chromosome 11 et le second sur le chromosome 6. Cependant, les déviations phénotypiques observées dans les IRCS suggèrent des incompatibilités moléculaires entre les fragments *spretus* et le fond génétique B6, autrement dit, des ruptures d'épistasies. De plus, les deux fragments contiennent 222 gènes, dont 40 sont des séquences annotées et 37 des séquences exprimées (de type Riken). Afin de cartographier le QTL impliqué dans le phénotype testiculaire il était indispensable de séparer les fragments *spretus* en créant des sous lignées congéniques.

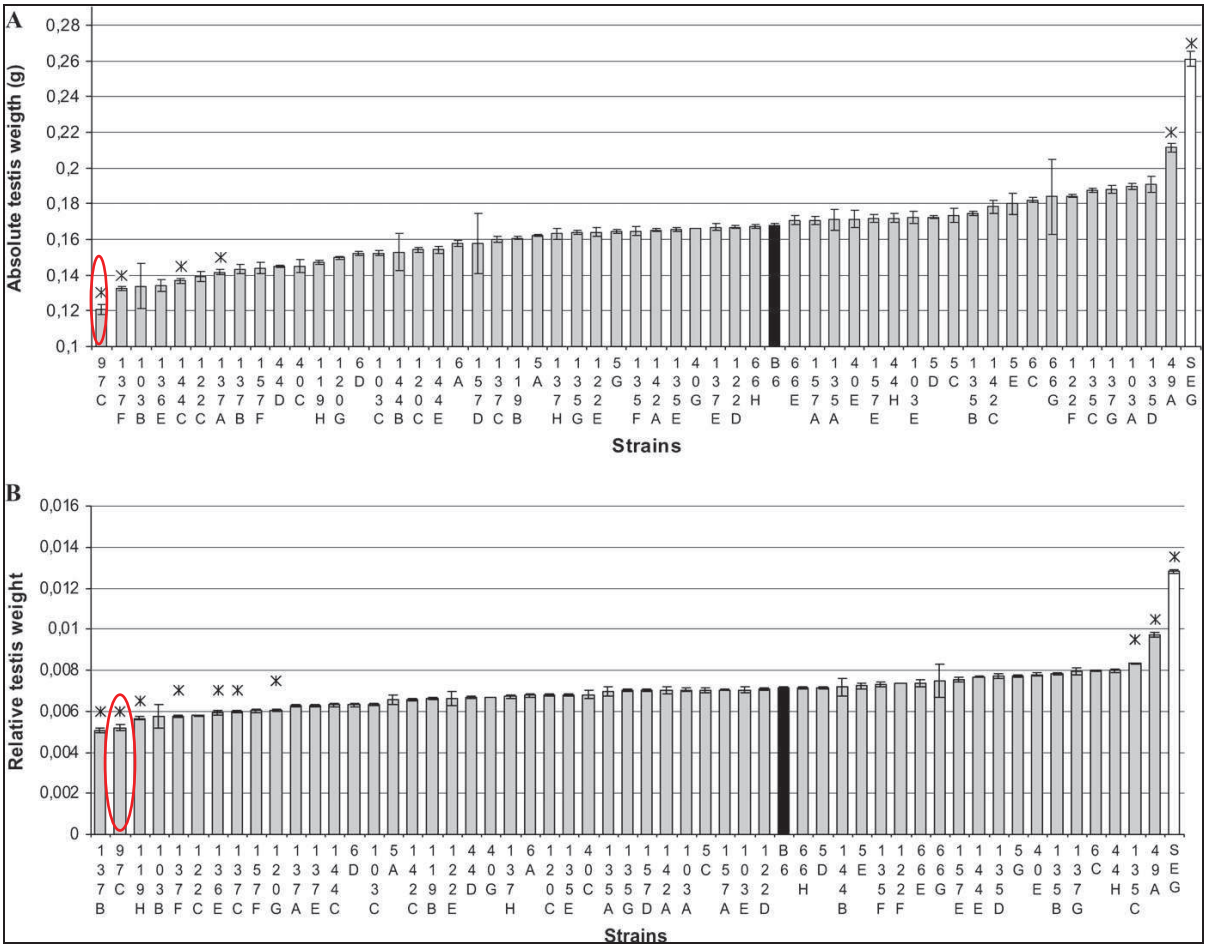


Figure 8: Poids testiculaire des IRCS et des lignées parentales (L'Hote et al., 2007). *Les valeurs moyennes (+/- SEM) absolues (A) et relatives (B) du poids des testicules sont présentés. Les astérisques (*) désignent les souches qui ont montré une différence significative dans les valeurs moyennes par rapport à B6.*

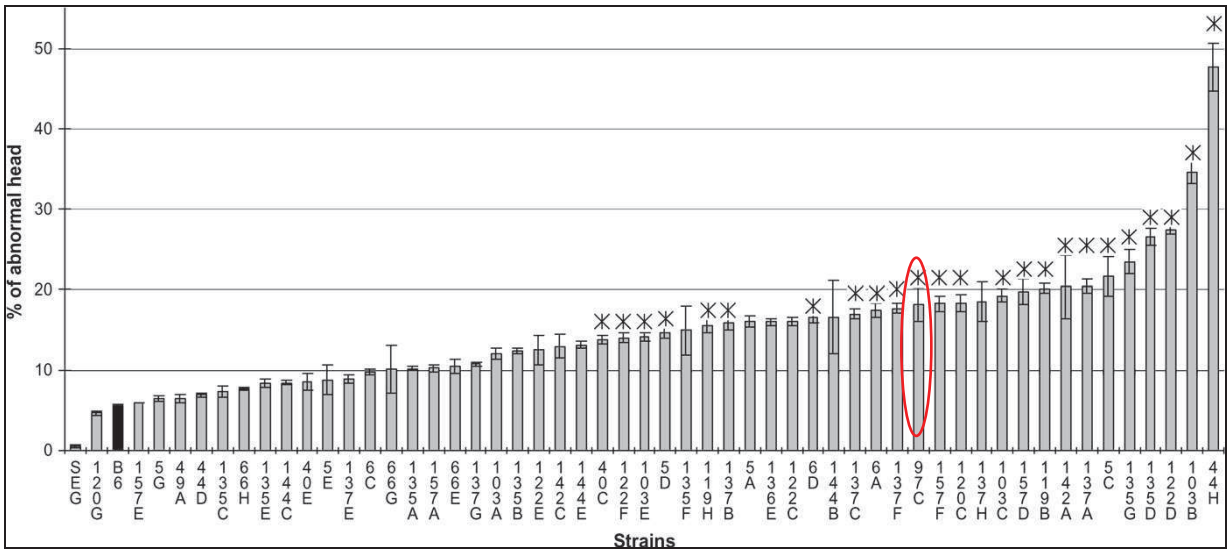


Figure 9: Fréquence d'anomalies de morphologie de la tête des spermatozoïdes. *Les valeurs moyennes (+/- SEM) sont présentés et les astérisques (*) distinguent les 24 souches qui ont montré une augmentation significative des valeurs moyennes par rapport à B6.*

Durant mon stage de M2 et ma première année de thèse, j'ai pu participer à la suite de ces travaux, à savoir la cartographie fine du QTL et l'identification du gène causal. Les résultats de ces travaux ont fait l'objet de ma première publication et seront présentés dans la partie résultats (L'Hôte et al., 2011).

B. Cartographie de QTL de fertilité femelle

Chez la souris (espèce polyovulatoire), la mort d'un ou plusieurs embryons in utero n'entraine pas la mort des autres embryons viables implantés

dans les cornes utérines. Les embryons morts subissent un processus involutif appelé résorption embryonnaire. Ce phénotype a été assimilé à des pathologies de la grossesse chez la femme, en particulier, l'existence de fausses-couches spontanées répétées.

Dans le but d'identifier de nouveaux gènes responsables des fausses-couches spontanées répétées chez la femme, Laissue et al. (2009) ont recherché des QTL de létalité embryonnaire dans les lignées IRCS.

Le phénotype de résorption embryonnaire a été évalué par échographie à haute fréquence après une période de mise au point en collaboration avec la plateforme d'imagerie du petit animal de l'institut Cochin. Les femelles IRCS et contrôles B6 sont croisées avec des mâles B6. Par l'observation du développement embryonnaire in vivo et en temps réel, les cinq étapes de la résorption embryonnaire ont été mises en évidence (Figure 10) (Laissue et al., 2009).

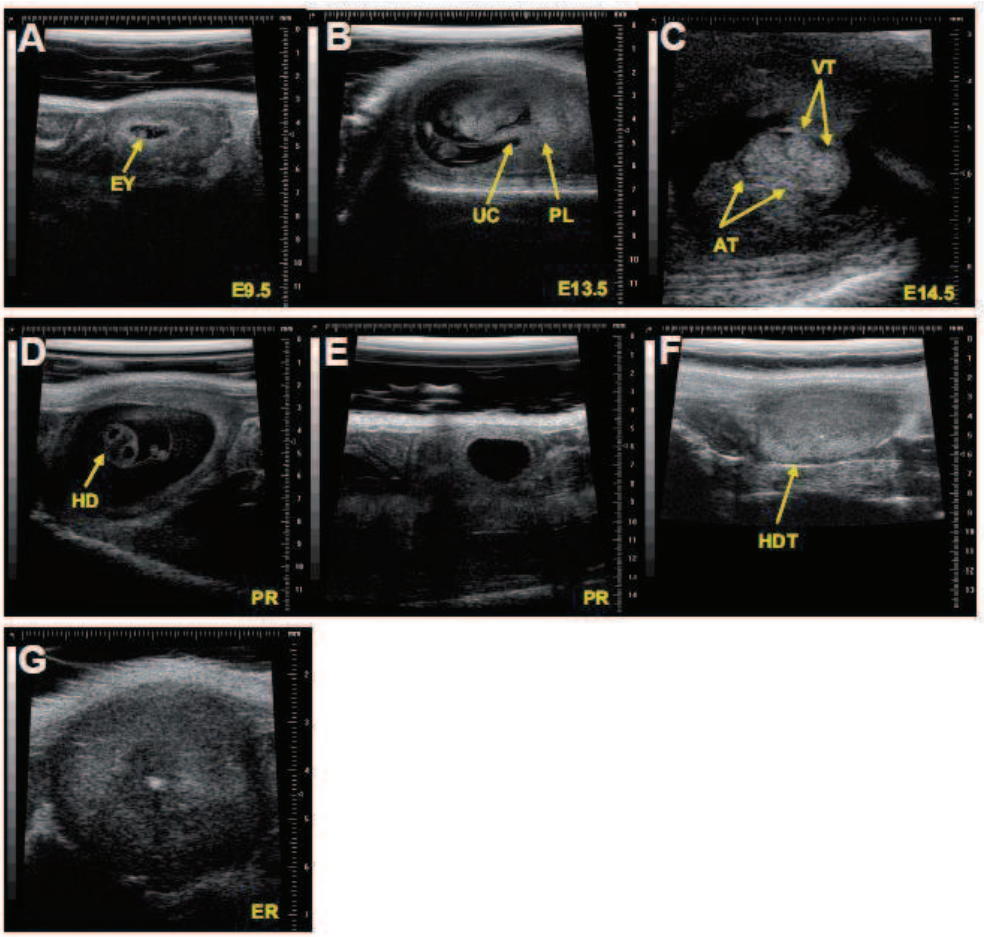


Figure 10 : Les 5 étapes de le résorption embryonnaire chez la souris (Laissue et al., 2009). *A : Embryon vivant à E9.5 ; B : Embryon vivant E13.5 ; C: Embryon vivant E14.5 ; D et E : Résorption progressive (PR) des structures embryonnaires ; F : Remplacement des structure par un tissu dense ; G : Point échogène caractéristique de la résorption embryonnaire (ER). EY, embryo; PL, placenta; UC, umbilical cord; VT, cardiac ventricles; AT, atrium; HD, head; HDT, high density tissue.*

L'évaluation du taux de létalité (nombre d'embryons morts par rapport au nombre d'embryons implantés), a notamment permis d'identifier une augmentation significative de la mort embryonnaire dans la lignée 66H qui possède 3 fragments *spretus* localisés respectivement sur les chromosomes 1, 13 et 18. Les fragments *spretus* ont été séparés dans des sous lignées congéniques

et leur phénotypage a permis de localiser 2 QTL, le premier Led1 dans la sous-lignée 66HMMU13 et le second Led2 dans la sous-lignée 66HMMU1 (Figure 11).

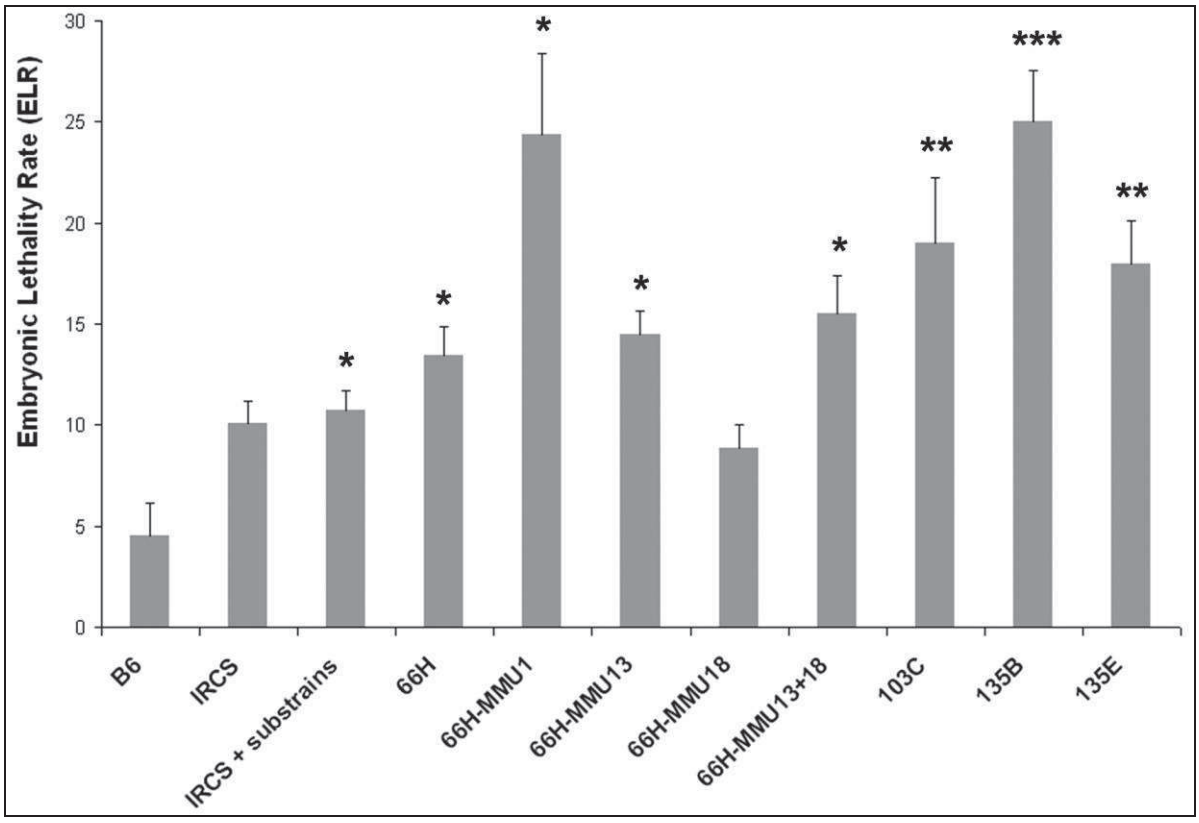


Figure 11 : Phénotypage des IRCS sur le taux de létalité embryonnaire (Laissue et al., 2009). *Le taux de létalité embryonnaire est défini comme la proportion d'embryons morts à n'importe quel stade du développement par rapport à la totalité des embryons implantés. Les résultats sont présentés comme le taux moyen d'embryons morts (+/- SEM).*

La cartographie du QTL Led1 dans la sous-lignée 66HMMU13 dont l'unique fragment *spretus* est d'environ 3 Mb, ainsi que des considérations transcriptomiques et bioinformatiques nous a permis de proposer qu'une

mutation du gène *Foxd1* serait responsable du phénotype de résorption embryonnaire. Les résultats ont ensuite été transposés à l'espèce humaine, en séquençant une collection de patientes et de contrôles, ce qui a permis d'associer des mutations de *FOXD1* aux fausses couches spontanées répétées.

Au cours de ma thèse j'ai collaboré à cette étude notamment en réalisant le phénotypage de lignées par échographie à haute fréquence (Article en préparation).

Cependant, l'essentiel de mon travail de thèse a consisté en la cartographie fine du QTL de mort embryonnaire Led2 dans la sous lignée 66HMMU1 dont l'unique fragment *spretus* était d'environ 30 Mb. Les résultats obtenus ont fait l'objet d'une publication (Vatin et al., 2012).

Objectif de la thèse

Ce travail de thèse a eu pour objectif l'identification et l'étude de nouveaux gènes responsables de défauts de fertilité mâle et femelle par l'exploitation des lignées IRCS. Dans la majorité des études présentées dans ce mémoire, la même approche a été réalisée : cartographie de QTL, clonage positionnel et, dans certains cas à la proposition du gène causal et son étude fonctionnelle.

Dans la perspective d'identifier des gènes impliqués dans la fertilité avec l'objectif de démontrer leur implication dans l'espèce humaine, une deuxième approche a été utilisée consistant à séquencer des gènes candidats chez des patients infertiles afin d'associer des polymorphismes à la pathologie étudiée.

Je me suis dans un premier temps intéressée au phénotype de défaut de spermatogenèse par l'étude de la lignée 97C présentant une réduction du poids testiculaire et une augmentation de la tératozoospermie par une dissection fine de la première vague spermatogénétique. Cette étude visait à renforcer l'hypothèse de l'implication du gène, Fidgetine like 1 que le clonage positionnel désignait comme le gène candidat le plus plausible. D'autre part, j'ai entrepris la caractérisation du phénotype des mâles d'une sous lignée de 66HMMU1 (Rc3)

qui présentent des anomalies de morphologie de la tête des spermatozoïdes associées à une diminution de leur pouvoir fécondant ceci afin de proposer des gènes candidats pertinents potentiellement responsables du phénotype. Dans les deux études des gènes candidats pertinents ont été identifiés dont le gène *Fignl1* et des études fonctionnelles ont été effectuées.

La seconde partie de mon travail de thèse a consisté à étudier le phénotype d'augmentation du taux de résorption embryonnaire dans les lignées IRCS 66HMMU1 et 66HMMU13 ainsi qu'à l'identification et l'étude de gènes candidats potentiellement responsables de cas de RSA. J'ai ainsi pu proposer 7 nouveaux gènes candidats, et des polymorphismes dans deux gènes (*FOXD1* et *ALPP*) ont été associés à des cas de RSA grâce au séquençage de patientes bien phénotypées (en collaboration avec le Pr Gris du CHU de Nîmes).

Résultats

Article 1

Fidgetin-Like1 Is a Strong Candidate for a Dynamic Impairment of Male Meiosis Leading to Reduced Testis Weight in Mice

(L'Hôte et al., 2011)

Résumé

Dans cette étude, nous avons réalisé le clonage positionnel d'un gène capable de moduler le poids des testicules par son action sur la méiose. Le QTL *Ltw1* (région *spretus* de 23 Mb sur MMU11) a été identifié dans la lignée IRCS 97C comme étant impliqué dans la diminution du poids testiculaire et une tératozoospermie modérée (20-30%) (L'Hote et al., 2007). Le faible poids des testicules est dû à des altérations cinétiques survenant lors de la première vague de la spermatogenèse (allongement anormal de la prophase des spermatocytes) associées à de l'apoptose des spermatocytes.

Dans le but de déterminer le gène responsable du phénotype observé, la région QTL Ltw1 a été affinée à une région de 5 Mb contenant seulement 11 gènes dont celui de la Fidgetin-like 1 (Fignl1) qui est apparu le gène le plus pertinent compte tenu du phénotype touchant la phase méiotique. En effet, parmi les 11 candidats de la région QTL, il est le seul gène qui est exprimé lors de la méiose au stade spermatocyte. De plus, il présente des polymorphismes non-synonymes entre les lignées parentales (*Mus spretus* et *Mus musculus*).

Le gène *Fignl1* code pour un membre de la famille des spastin/katanins, impliquées dans l'interaction avec le fuseau de microtubules et potentiellement donc dans la séparation des chromosomes au cours de la méiose et de la mitose.

Il semble donc être un acteur important de la cinétique de la méiose mâle et un nouveau candidat potentiel pour expliquer des cas de stérilité masculine.

OPEN ACCESS Freely available online

PLoS one

Fidgetin-Like1 Is a Strong Candidate for a Dynamic Impairment of Male Meiosis Leading to Reduced Testis Weight in Mice

David L'Hôte[1,2], Magalie Vatin[3,4], Jana Auer[3,4], Johan Castille[6], Bruno Passet[6], Xavier Montagutelli[5], Catherine Serres[3,4,5], Daniel Vaiman[3,4,5]

1 CNRS UMR 7592, Institut Jacques Monod, Equipe Génétique et Génomique du Développement Gonadique, Paris, France, 2 Université Paris Diderot-Paris VII, Paris, France, 3 U1016 Département de Génétique et Développement, Institut Cochin, INSERM, Paris, France, 4 Université Paris Descartes, Paris, France, 5 Unité de Génétique des Mammifères, Institut Pasteur, Paris, France, 6 Département de Génétique des Animaux, INRA, Jouy-en-Josas, France

Abstract

Background: In a previous work, using an interspecific recombinant congenic mouse model, we reported a genomic region of 23 Mb on mouse chromosome 11 implicated in testis weight decrease and moderate teratozoospermia (~20–30%), a Quantitative Trait Locus (QTL) called Ltw1. The objective of the present study is to identify the gene underlying this phenotype.

Results: In the present study, we refined the QTL position to a 5 Mb fragment encompassing only 11 genes. We showed that the low testis weight phenotype was due to kinetic alterations occurring during the first wave of the spermatogenesis where we could point out to an abnormal lengthening of spermatocyte prophase. We identify Fidgetin-like 1 (Fignl1) as the gene underlying the phenotype, since it fulfilled both the physiological and molecular characteristics required. Indeed, amongst the 11 positional candidates it is the only gene that is expressed during meiosis at the spermatocyte stage, and that presents with non-synonymous coding variations differentiating the two mouse strains at the origin of the cross.

Conclusions: This work prompted us to propose Fignl1 as a novel actor in mammal's male meiosis dynamics which has fundamental interest. Besides, this gene is a new potential candidate for human infertilities caused by teratozoospermia and blockades of spermatogenesis. In addition this study demonstrates that interspecific models may be useful for understanding complex quantitative traits.

Citation: L'Hôte D, Vatin M, Auer J, Castille J, Passet B, et al. (2011) Fidgetin-Like1 Is a Strong Candidate for a Dynamic Impairment of Male Meiosis Leading to Reduced Testis Weight in Mice. PLoS ONE 6(11): e27582. doi:10.1371/journal.pone.0027582

Editor: Alexander J. Travis, Cornell University College of Veterinary Medicine, United States of America

Received July 14, 2011; **Accepted** October 19, 2011; **Published** November 16, 2011

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vaiman@cochin.inserm.fr

☯ These authors contributed equally to this work.

Introduction

Fertility is an important concern both in human medicine, where 10–15% of the couples call for the services of assisted reproductive technologies, and for various domestic species of economic interest, such as dairy cows where a striking drop in fertility has been observed in recent years [1,2]. In humans, the genetic basis of male infertilities is far from being completely elucidated and is mostly explained by deletions of the AZF region of the Y chromosome [3,4]. However, such infertilities affect no more than 10% of the male patients, stressing the need of identifying new actors responsible for these disorders. Hundreds of genes involved in gametogenesis in a broad sense have been identified by gene-invalidation approaches in mice [5]. But, while these experiments lead to a complete abolishment of gene expression, more quantitative approaches are of interest to identify new fertility genetic determinants. Up to now, several loci implicated in the quantitative regulation of both male and female fertility traits have been mapped in the mouse genome, thanks to crosses between closely related strains. However the genetic bases underlying these QTL (Quantitative Trait Loci) have generally not been discovered so far [6].

Different types of mouse models have been developed for accelerating positional cloning procedures, and are well-suited for identifying gene(s) governing quantitatively fertility traits. Amongst the common mapping tools existing in plant and animals, panels of inter-specific recombinant congenic strains (IRCS) have been developed. Such panels thanks to their interspecific origin, exhibit a much higher genetic diversity than intra-specific panels, thus enhancing the contrasts between phenotypes [7]. This is the case of a panel of 53 recombinant congenic mouse strains harboring a small quantity of *Mus spretus* genomic fragments dispersed in a *Mus musculus domesticus* background. This panel of IRCS has been developed at the Pasteur Institute (Paris, France) from an original cross between the SEG/Pas strain (SEG) of *Mus spretus* origin and the C57BL6/J (B6) *Mus musculus* strain [7,8]. On average, each IRCS genome is composed of ~1.5% of *Mus spretus* genome, distributed in 1 to 8 fragments in a C57B6/J context. *Mus spretus*

and *Mus musculus* diverged about 2 million years ago, this divergence leading to the existence of a nucleotide substitution roughly every 80 bp, close to that existing between humans and chimps [9]. This important diversity and the weak proportion of *Mus spretus* genome dispersed in B6 background make it possible to rapidly map QTL in the centimorgan range [10].

In a previous study, we reported on the mapping of a low testis weight QTL (*Tstw1*) mapping on mouse chromosome 11 (between 3.7 Mb and 26.5 Mb) following the phenotypic analysis of the 97 C strain [10]. This strain exhibits a significant reduction in absolute and relative testis weight compared to the B6 control, associated with a reduction in the seminiferous tubules' diameter. We then strived to identify the gene(s) responsible for this low testicular weight phenotype. For this, a double approach was undertaken aiming (1) at refining the location of the critical region by creation of recombinant events inside the *spretus* fragment of the 97 C mice, in parallel with (2) a thorough phenotypic analysis of the testis in order to pinpoint and characterize as precisely as possible the features differentiating 97 C and B6 mice. A combination of these cartographic and phenotypic approaches, together with a careful exclusion of most the genes present in the minimal remaining fragment, called attention to the *Fidgetin-like 1* gene (*Fign1*) as the best candidate for explaining the 97 C testis phenotype. Here, we propose that *Fidgetin-like 1* polymorphisms differentiating the two mouse strains are responsible for the meiosis phenotype observed. In this study, we present arguments suggesting that the role of this gene is to control male meiosis dynamic.

Results and Discussion

1) Fine phenotyping of 97 C testes reveals an impaired progression of the germ cells through the first wave of spermatogenesis

We extended our preliminary observations of the 97 C phenotype [10] on an increased number of animals ($n = 19$) for testis histology and seminal vesicle secretion. This analysis did not reveal any obvious abnormality concerning spermatogenesis in 97 C mice and confirmed the observation of a clear differences in testis weight between 97 C and B6 adult mice (0.12 g instead of 0.18 g, corresponding to a difference of ~30%, Figure 1a). Animals harbored a normal epididymal sperm reserve, associated to a teratozoospermia about 3 to 4 fold more elevated than in the B6 parent (20% vs 6%). Despite this recurrent defect affecting sperm head shape, 97 C animals were fertile, probably due to the high percentage of normal spermatozoa remaining in this strain. The sexual behavior of 97 C males was not different from the B6 parent, indeed, vaginal plugs were observed in mating experiment with B6 female, and 97 C male sired normal sized litters.

Since our previous characterization of the 97 C testicular phenotype was limited to adult mice [10], we examined the postnatal testis development in 97 C mice in order to detect putative early defects at the origin of the adult phenotype. Testes were analyzed at various time points from birth until two months (Figure 1). The testis weight evolution in 97 C mice was apparently normal from 9 to 16 days post partum (DPP) but was delayed from 21 DPP. This difference became statistically significant from 28 DPP and remained significant in older mice. 97 C testis weight reached a plateau at 6 weeks, exhibiting a value about 30% lower than B6's in adult mice (Figure 1a). In parallel with testis weight, the mean diameter of seminiferous tubules in 97 C mice was significantly smaller than this of B6 from 21 DPP (~7%) and this difference reached 25–30% thereafter (Figure 1b). Since this time period corresponds to the onset of the first wave of spermatogenesis, we examined in detail histological sections during this period (figure 2) and we quantified the seminiferous tubule sections according to the stage of the spermatogenesis cycle (figure 3). No obvious difference was observed between 97 C and the B6 testis histology at 9 DPP and 14 DPP (figure 2 a to d). At 14 DPP we observed tubules containing spermatocytes in similar frequency in 97 C and B6 testes (90.5±4.61% vs 85.0±3.46%, respectively), at leptotene, zygotene and pachytene stages (figure 3a) suggesting that the proliferative phase of spermatogenesis and the entry in meiosis occur normally. The first sign of dysfunction in 97 C spermatogenesis was detected after the third week after birth. At 24 DPP, we observed in B6 testes a majority of sections with round spermatids (figure 2f and figure 3b) whereas in 97 C testes, the tubules containing this post meiotic stage were significantly less frequent (65.6±6.56% vs 37.37±9.08% respectively, $P<0.001$ (figure 3b)). We could still detect in 97 C testes a majority of tubules with spermatocytes as the most advanced stage, notably early (first mid) pachytene stage, which were less frequent in B6 testes (56.2±10.35% vs 20.07±5.30% respectively, $P<0.001$). At 28 DPP, tubule sections of 97 C testes containing spermatocytes were still observed, at the same frequency than those with round spermatids (45.1±2.69% vs 47.7±3.11% respectively) (figure 2g and figure 3c). At this same period in B6 testes, spermiogenesis concerned more than 90% of the germ cells and was already advanced to the elongated and condensated spermatid stage (13–15 step) (figure 2h and figure 3c). At 35 DPP, half of the observed sections of 97 C testis still contained both delayed meiotic cells and round spermatids (step 1–7), stages nearly absent in B6 testis (figure 3d) The release of testicular spermatozoa in the tubule lumen was visible at 35 DPP for B6 and 42 DPP for 97 C (figure 2j and h). Consistently with this time-lag in spermatogenesis differentiating the two strains, the central lumen of tubules in 97 C testis was not formed before 4 weeks, whereas in B6 testis the lumen was opened from the third week.

The examination of epididymis sections over the same period of time (2 to 6 weeks) highlights a one week delay in the apparition of spermatozoa in the epididymal duct of 97 C compared to B6 mice (week 6 versus week 5, respectively) (Figure S1).

Examination of seminiferous tubule sections of SEG strain at 14, 24 and 28 DPP did not reveal difference in the time course of spermatogenesis compared with B6 mice (data not shown).

To determine whether the delay in the spermatogenesis development of 97 C mice was accompanied with apoptosis, we performed a TUNEL assay on adult testes (in post-natal rodent testis, cell apoptosis naturally occurs in spermatogonia and spermatocytes during the first wave of spermatogenesis, consequently we decided to assess apoptosis during adulthood in mature testis). The frequency of positive tubules (e.g. section where at least one apoptotic cell was detected) was significantly higher in 97 C testis when compared to B6 (0.30 and 0.18 respectively; chi square test $P<0.007$). Considering only positive sections, the mean number of apoptotic cells per tubule was significantly higher in 97 C testis than in B6's (3.98±3.15 and 2.40±1.54, respectively, $P<6.7\ 10^{-5}$). This averaged rate masked a variable repartition among tubules, with apoptotic cells often observed in clusters in 97 C sections whereas they were rather scattered in B6 sections (Figure 4). The apoptotic cells in 97 C testes were mainly spermatocytes at mid-pachytene stage (see Figure 4). This increase in apoptotic cell number in 97 C, although clearly reflecting a dysfunction, is not sufficient to significantly reduce the sperm production of 97 C adult mice [10].

Taken together, these results show that in the 97 C testis, spermatogenesis progresses normally up to the pachytene spermatocyte stage, but a functional impairment in the first half of pachytene stages (still visible at 3 and 4 week) slow down the

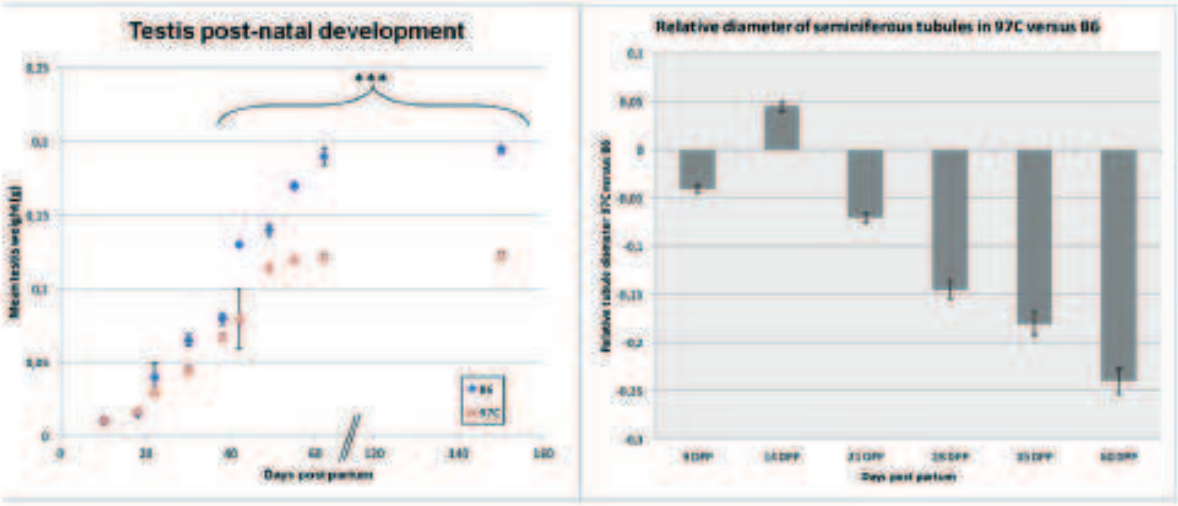


Figure 1. Testis development analysis of 97 C mice. The testis weight and the seminiferous tubule diameter were quantified in function of the age in B6 and 97 C RCS mice. (a) The testis weight of B6 (▲) and 97 C (◆) mice was expressed as a mean ± SEM of 3 to 4 animals per strain. *** indicates a significant difference with $p<0.0001$ between 97 C and B6 mice. (b) Tubule diameters of B6 and 97 C mice were measured on 3–4 testes per strain on an average of 40 tubules. The results represent the ratio of the difference between 97 C and B6 mean values relative to B6 value. From 21 days post partum, the difference is significant ($p<0.001$).
doi:10.1371/journal.pone.0027582.g001

passage to late pachytene resulting in a lengthening of the meiosis duration. This defect, detected during the first wave of spermatogenesis becomes less evident in adult testis where successive spermatogenetic waves are overlapping. However the rate of apoptotic spermatocytes higher in adult 97 C testis than in B6 'one arrests for such a dysfunction. Thus, this early defect is never overcome and is presumably at the origin of the adult phenotype since the first wave of spermatogenesis and testicular growth are two correlated events. Such an association between meiosis delay and reduced adult testis weight has been observed, for example, in mouse mutants for TR4 genes [11]. In sum two possible effects could lead to the reduced testis weight: a slowing down of the passage from mitosis to meiosis of the spermatogonies, or a slowing down of the meiosis possibly inducing the observed increased apoptosis of germ cells in the seminiferous tubules, a hypothesis that we privilege to explain the phenotype.

As previously mentioned [10], 97 C males presented also with a seminal vesicle phenotype showing a translucent appearance instead of the normal off-white color visible in B6. The secretion of theses seminal vesicles was further examined by a 1D protein gel stained by Coomassie blue. No obvious difference was visible between B6 and 97 C.

2) Fine mapping of the low testis weight QTL (Ltw1) of 97 C line

In a previous study, we have mapped a QTL of low testis weight (*Ltw1*) on a 23 Mb region of chromosome 11, thanks to the 97 C line supposed to harbor a unique *genus* fragment on MMU11. The localization of this QTL was confirmed by the analysis of a F2 population [10]. On the occasion of an expressional study of testis genes of 97 C, we discovered an additional *genus* fragment of approximately 9 Mb on MMU6 in this same strain, this fragment having been detected by its specific expression profile, compared to the B6 background [12]. In order to unambiguously map the *Ltw1* QTL, we created two congenic sub-strains harboring one *genus* fragment, either on MMU6 or on MMU11 (called 97Crc6 and 97Crc11 respectively). Phenotypic analysis of these new strains showed that 97Crc6 mice exhibited only the seminal vesicle phenotype described above (abnormal liquid coloration) and that their mean testis weight was not significantly different from B6 (figure 5). As expected, the low testis weight phenotype segregated with the MMU11 *genus* fragment harbored by the 97Crc11 strain (figure 5). We confirmed that 97Crc11mice exhibited the same abnormality of spermatogenesis that 97 C mice.

In order to identify the genetic determinant of the *Ltw1* QTL, we first refined the MMU11 *genus* fragment spreading over 23 Mb (3,539797 to 26,483225 bp) to a smaller QTL region of 16 Mb (3,539797 to 19,470000 bp) by genotyping microsatellites at the fragment boundaries. This region still encompasses 115 genes (either annotated or expressed sequences). For improving the mapping resolution, we backcrossed 97Crc11 mice with B6 parents in order to generate recombination events inside the MMU11 *genus* DNA segment and then we derived from them homozygous sub-congenic strains. Three informative recombinant strains were generated at the homozygous state, 97Crc11a (14.7 Mbp–19.5 Mbp), 97Crc11b (13.4 Mbp–19.5 Mbp), 97Crc11c (9.6 Mbp–19.5 Mbp) (figure 6) and analyzed for the testicular phenotype. Only 97Crc11c exhibited a significantly smaller testis weight than that of B6 (figure 6) with a high level of teratozoospermia (35% of abnormal spermatozoa in average). The overlapping of the *genus* segments of the 3 recombinant lines narrowed down the minimal interval of the *Ltw1* QTL to a 5 Mb *genus* region located between microsatellites D11MIT259 and D11MIT148 (9.6–14.7 Mb). The reduced *Ltw1* QTL region of interest contains only 11 annotated genes according to the NCBI database (Table 1).

3) Fidgetin-like1 is a pertinent positional and functional candidate gene susceptible to affect spermatogenesis in 97 C mice

To point out a pertinent candidate gene we choose to apply several screens considering expressional and functional data provided by experimental results, databases, and literature reports.

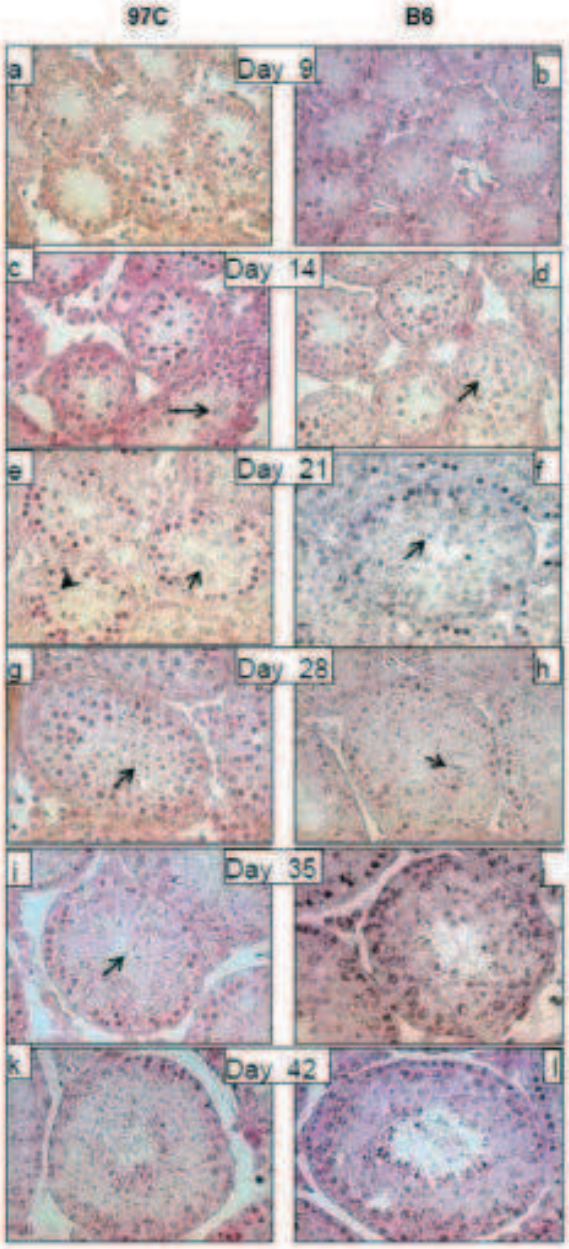


Figure 2. Impaired progression in the first wave of spermatogenesis in juvenile 97 C and B6 parental mice. A representative picture of seminiferous tubule histology of testes from 97 C (a,c,e,g,i,k) and B6 (b,d,f,h,j,l) mice at post natal day 9, 14, 21, 28, 35 and 42 is shown (objective ×60). At day 9 and 14, no difference was found between B6 and 97 C tubule cell population. A germ cell progression up to pachytene spermatocytes (arrows) was observable in both 97 C (c) and B6 (d) testes. At day 21, meiosis was completed and round spermatids (arrow in f) appeared in the tubule of B6, while spermatogenesis of 97 C was still in the meiotic phase with mostly germ cells at pachytene spermatocyte (arrow in e) or metaphase stage (arrow head in e). In addition, at this age, the lumen of the tubules was visible only in B6 mice. At day 28, spermatogenesis normally progressed in B6 up to the elongated spermatid stage (arrows in h) but displayed a one week delay in 97 C testes, where round spermatids (arrow in g) were the most advanced stage of germ cells present in the tubules. Elongated spermatids were not observed before the fifth week in post-natal tests of 97 C (i). Figures of spermiation seen at day 35 in control testes (h) signed the end of the first wave of spermatogenesis in B6 mice, which was only attained at the sixth week in 97 C mice (k). From this time, testis histology of the two strain was comparable (k and l) but a smaller tubule diameter was noticeable for the 97 C testis.
doi:10.1371/journal.pone.0027582.g002

In a previous study, we performed a whole genome expressional analysis on B6 and 97 C testis [12]. We took profit of these data to quantify the expressional level of the 11 genes presents in the 5 Mb QTL region in B6 and 97 C testes (Table 1). Considering only transcripts with a fluorescence level >200 (arbitrary threshold for significant expression) on the Nimblegen arrays used, we noted that four genes *Vac2*, *LOC634636*, *Ikzf1* and *Gbl* are expressed at a low or undetectable level in the B6 testis whereas four other genes, *Gtl10*, *Zfp6*, *4930415F15Rik* and *Figla1* are expressed at a high level (1,400 to 16,600 fluorescence unit, average fluorescence in the array ~2300). For the rest of the study and taking into account its expression profile, we decided to call *4930415F15Rik* 'Pmst1', for Post-Meiotic Spermatogenesis 1). These experimental results were consistent with the bioinformatic data (SymAtlas and Germonline data bases) reporting either no expression (for *Pmst1*, *Vac2*, *Ddc*, *Gbl*, *Ikzf1*) or a moderate (for *Gtl10* to strong (for *Figla1*, *Zfp6* and *Pmst1*) expression in the testis compared to other tested tissues.

The expression level of the four genes significantly expressed in the testis (*Gtl10*, *Figla1*, *Zfp6* and *Pmst1*) was not different between 97 C and B6 mice as shown by the expression ratio 97 C/B6 which ranged from 0.8 to 1 (Table 1). Given this absence of deregulation, we attempted to identify sequence polymorphisms differentiating the *spretus* alleles from the *musculus* alleles of these genes. Systematic sequencing indicated that *spretus* alleles of *Zfp6* and *Gtl10*, in SEG/Pas and 97 C strains, present only with synonymous variations by comparison with *musculus* alleles present in the B6 strain. By contrast, non synonymous polymorphisms were found in the two other genes, with one variation in *Pmst1* (p. glu100val) and 9 differences in *Figla1* (figure 7a). Moreover, we found that 97 C and SEG testes expressed an additional truncated isoform of *Figla1* corresponding to alternative splicing of Figla1 mRNA, as revealed by sequencing (figure 7b).

In our search of the candidate gene, we could reasonably exclude *Vac2*, *LOC634636*, *Ikzf1*, *Gbl*, *Zfp6* and *Gtl10* genes estimating that their implication in the phenotype was improbable, due to their low level of expression in the testis (*Vac2*, *LOC634636*, *Ikzf1* and *Gbl*), or their absence of expressional deregulation in 97 C testis as well as their absence of non-synonymous polymorphisms (*Zfp6* and *Gtl10*). This exclusion was strengthened by the results of gene invalidations for *Zfp6*, *Gtl10* and *Ikzf1* which give rise to phenotypes far from that of 97 C mice. Indeed the knock-out of *Zfp6*, a gene principally expressed in post-meiotic germ cells, induces an aberrant acrosome biogenesis and dysmorphic spermatozoa, dissimilar to the characteristics observed in 97 C [13]. The disruption of *Gtl10*, which encodes a growth factor receptor-binding protein, leads to a disproportionate overgrowth of some organs (placenta, muscle and pancreas) whereas the overexpression of some isoforms of the encoded protein results in growth suppression [14] a phenotype that was not observed for 97 C mice. At last, the mouse mutant for *Ikzf1* gene displayed an abnormality in the blood lineage with stem cells differentiating exclusively into erythroid and myeloid cells without providing lymphocyte cells [15,16].

Thus, only two relevant potential candidate genes remained: *Pmst1* and *Figla1*. *Pmst* which is highly and specifically expressed in

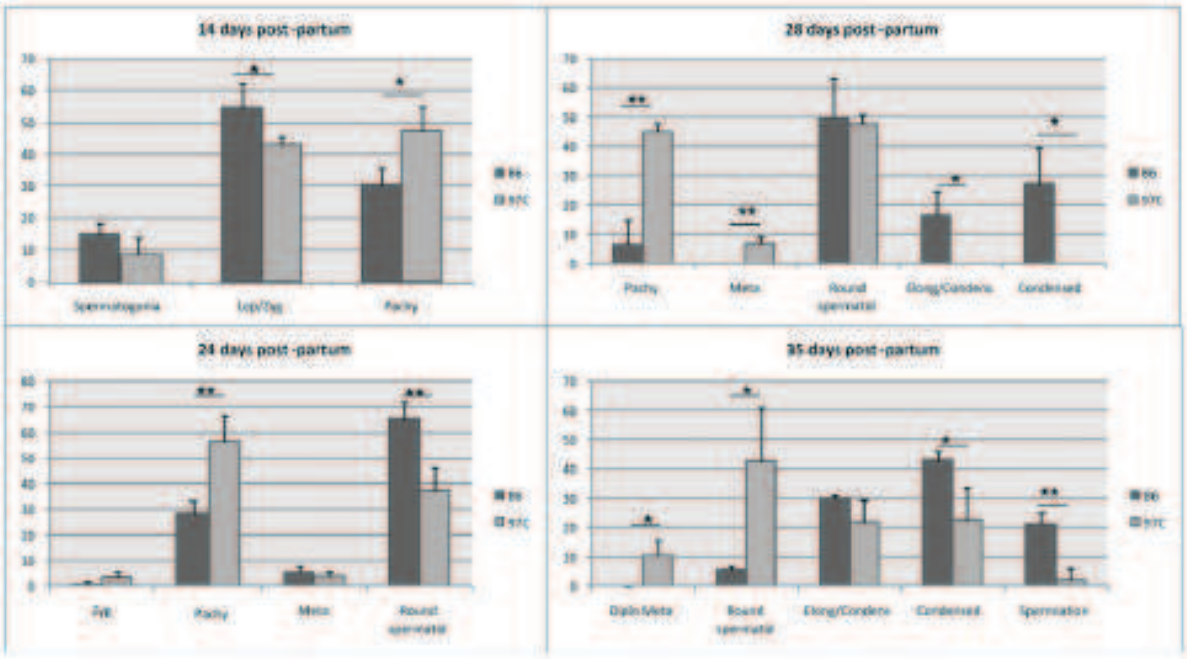


Figure 3. Distribution of the different stages of the spermatogenesis during the establishment of the first wave from 14 days post partum to 35 days post partum in testes of B6 (black bars) and 97 C (grey bars) mice. The results represent mean ± SD of the frequency of each stages observed in the testes of 3 or 4 males for each strain (with 50 to 130 tubule sections observed by male at each age). Abbreviation: Lep – leptotene; Zyg – zygotene; Pachy – pachytene; Diplo – diplotene; Meta – metaphase; Round spermatid corresponds to the steps 1 to 7 of the spermiogenesis; Elong/condens – elongating and condensing spermatid of steps 8 to 12 of the spermiogenesis; Condensed – condensed spermatid of steps 13 to16 of spermiogenesis. Different between B6 and 97 C at:* $p<0.05$; ** $p<0.01$; *** $p<0.001$ with student *t* test.
doi:10.1371/journal.pone.0027582.g003

the mouse testis, presented with a post-meiotic expression at the spermatid stage. Indeed, during the first wave of spermatogenesis analyzed by RT-PCR, *Pxt1* was undetectable until 20 DPP, a period where seminiferous tubules were principally populated by spermatogonies and spermatocytes and became detectable only afterwards (22–30 DPP) when the spermatids are present in addition to the precedent cells (Figure S2). We confirmed this expression pattern at the protein level by western blots using an antibody designed and characterized for the present study (Eurogentec ™) (data not shown). Therefore it is highly unlikely that a gene strictly expressed at a post-meiotic spermatid stage could affect a meiotic process. During this experimental work, we showed that the LOC432534 annotated gene is in fact a part of a *Pxt1* transcript. Indeed we succeeded in amplifying by RT-PCR on mouse testis cDNA the whole transcript spanning from the AUG codon of LOC432534 to the stop codon of *Pxt1*. We showed by RT-PCR that this new isoform also exists on testis cDNA from human, dog, cat and chicken. This new isoform 2 of *Pxt1* displayed the same expressional post meiotic pattern than the isoform 1 (NM_028669) (data not shown).

Contrarily to *Pxt1*, *fidgetin like 1* (*Figl1*) is reported to be expressed in pachytene spermatocytes and to a lesser extent in spermatogonia. *Figl1* encodes a protein belonging to the group of ATPases Associated with diverse cellular Activities (AAA) [17] in the subgroup of meiotic proteins which also comprises *fidgetin*, *fidgetin-Like 2*, *katanin*, *spastin*, and *SKD1* genes [18]. We realized an immunolabeling of adult testis sections with an anti-FIGNL1 antibody. Histological analysis of B6, 97 C and SEG testis sections revealed a same localization of Fignl1 for the three strains, in spermatocyte cells (figure 8). A perinuclear staining was visible in the cytoplasm of pachytene spermatocytes (figures 8 a, b, c) and a strong immunoreactivity was also observed in metaphase spermatocytes (figure 8 d, e, f). By contrast no signal was found in spermatid cells (large arrows in enlarged figures 8 a, b, c). These experiments revealed a quite striking exclusion of labeling of the spindle area in meiotic metaphase in B6, while in 97 C, and to a lesser extent in SEG, some metaphasic spindle area appeared labeled with anti FIGNL1. In agreement with the testis transcriptome data showing a same expression level for B6 and 97 C *Figl1* (Table 1), a Western blot of adult testes from B6, SEG and 97 C did not reveal differences in expression levels of Fignl1 protein detected at the 74 kDa expected size (Figure 9a). An analysis by WB of the presence of Fignl1 during the first spermatogenetic wave in B6 testis showed that it was still undetectable at 10 DPP, i.e. at a moment close to the entry in male meiosis. It was detected at 16 PDD and persisted afterward, in concordance with its expression at the spermatocyte stage (figure 9b).

To better characterize Fignl1, we performed a two-dimensional electrophoresis/WB of B6 and 97 C testis extract (Figure 9c). The profile of 2D-migration of Fignl1 according to the isoelectric point (pI) and the molecular mass (MM) was the same for 97 C and B6 testis, i.e. the protein migrated as two principal dots at 76 kDa and 72 kDa for a same 5.86 pI. A third faint but detectable spot migrated at 76 kDa and 5.76 pI. These values are slightly different from the theoretical values (74 kDa MM and 6.04 pI), signifying moderated posttranslational modifications. The pI values slightly more acidic compared to the theoretical value (difference of 0.18

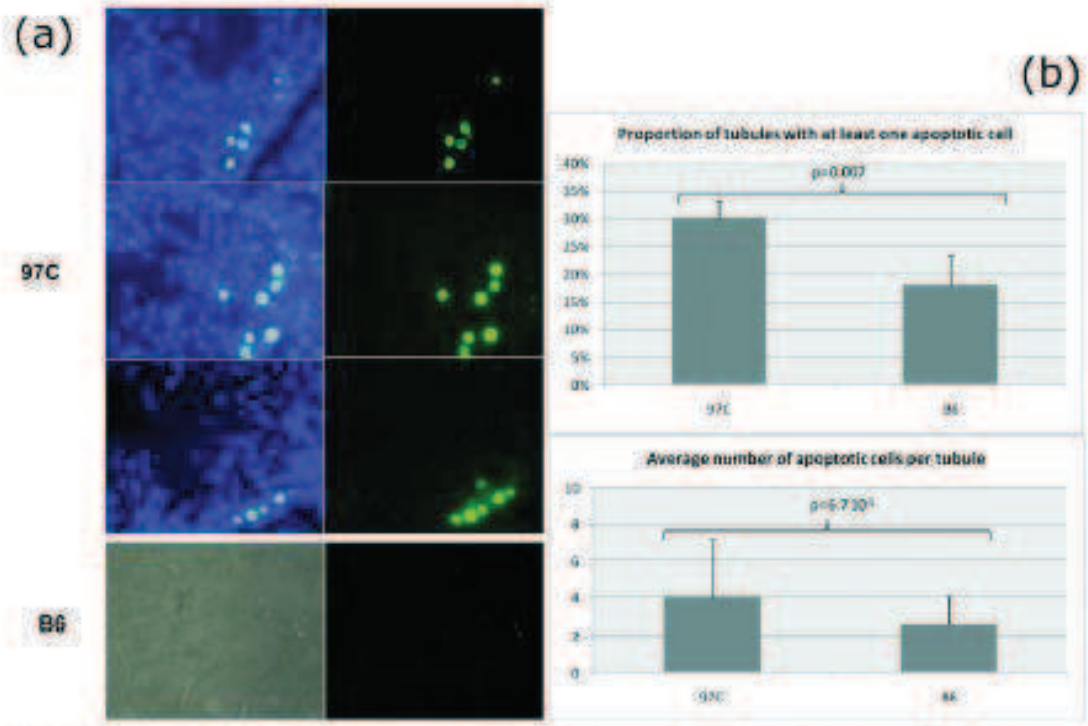


Figure 4. Apoptosis in the seminiferous tubules of adult 97 C mice. The apoptotic cells were studied for their localization (a) and their number (b) in B6 and 97 C testes using the TUNEL assay. (a) DNA strand breaks of the apoptotic cells were revealed by fluorescein-dUTP (green fluorescence) and the sections were counterstained by DAPI (blue fluorescence). Clusters of TUNEL positive cells were observed in tubule sections of 97 C testes, principally spermatocytes at the stages VI–VII of the spermatogenesis (objective ×100). In B6 testis sections, the apoptotic cells were scattered in some tubules (left transmission light, right epifluorescence; objective ×20). (b) Percentage of sections containing at least one apoptotic cell was significantly higher in 97 C testis when compared to B6 (30% and 18% respectively; chi square test $p<0.007$). In these positive tubules, the number of apoptotic cells per tubule (mean ± SD) was significantly higher in 97 C testis than in B6's (3.98±3.15 and 2.48±1.54, respectively; $p<6.7\ 10^{-9}$).

doi:10.1371/journal.pone.0027582.g004

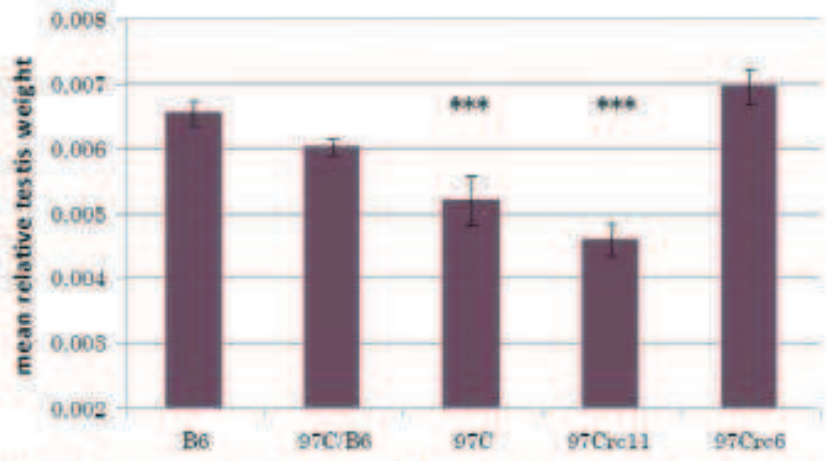


Figure 5. Relative Testis weight of parental B6 and IRCS 97Crc6, 97Crc11, 97 C. 97 C encompasses two chromosome segments of spretus origin on chromosomes 6 and 11. The 97Crc6 and 97Crc11 sub-recombinant were generated to separate the two fragments. Clearly the small testis weight phenotype segregates with chromosome 11. Results are expressed as mean ± SEM. Asterisks (***) denote strains that showed a significant difference when compared to B6 ($P<5.10^{-7}$).

doi:10.1371/journal.pone.0027582.g005

Figure 6. Fine mapping of the low testis weight QTL in inter-recombinant strains generated for this study from 97Crc11. Left panel: genotypes of the different recombinant lines (97Crc11a, b, c, noted 97Ca, b and c in the figure), parent B6, 97 C and heterozygote 97 C/B6. Black regions correspond to B6 background, white regions to homozygous *spretus* fragments and grey regions to heterozygous *spretus* fragments on chromosome 11. Marker positions are given in megabase pairs. Right panel: Histogram of the testis weight relative to body weight expressed as mean ± SEM for a minimum of 5 animals. *** denotes a significant difference when compared to B6 ($p<10^{-3}$). By analyzing the fragment/phenotype segregation, it could be concluded that since 97 C and 97Crc11c exhibit the phenotype of low testis weight, the *Ltw1* QTL is located in a fragment of 5 Mb or less on MMU11 (hachured region between 9.6 Mb to 14.7 Mb).
doi:10.1371/journal.pone.0027582.g006

and 0.20 unit) could correspond to two phosphorylated sites and one acetylated site of the Fignl1 protein as found by high throughput experiments [19,20,21]. In sum, Fignl1^{spretus} protein expressed in 97 C testis does not appear biochemically different from Fignl1^{domus}, protein expressed in B6.

These overall results that describe for the first time the presence and the subcellular localization of Fignl1 in mammalian testis indicate that Fignl1 is mainly localized in meiotic germ cells, the spermatocytes, with a subtle anomaly of its localization in 97 C testis. Since spermatocytes are functionally impaired in 97 C testes, our results on Fignl1 together with the exclusion of all the other genes present in the minimal fragment strengthen the implication of Fignl1 in the *Ltw1* QTL.

In the literature, three studies give consensual insights on *Fignl1* function reporting action on proliferative/differentiation processes. Very importantly, the loss of *Rgl1* function by injection of anti-*Fignl1* siRNA in *C. elegans* [22] leads to the accumulation of mitotic nuclei inside the proliferative zone of the gonad, preventing the passage of the cells to meiosis. Another study showed that in transiently transfected mouse osteoblast cell model (MC3T3-E1) Fignl1 inhibits osteoblast proliferation and stimulates their differentiation [23]. The third study, using a silencing approach in proliferating MIN6 cells (derived from pancreatic β-cells), suggested that Fignl1 was implicated in adequate accomplishment of cell cycle by controlling cell survival [24]. Thus, if Fignl1 function is impaired in the 97 C testis, it could lead to a slowing down of germ cell differentiation with, in consequence, an increase of apoptotic cells (it is natural to think that defects in the normal progression of the meiotic cell cycle must to end with apoptosis to avoid formation of defective gametes). The extension of meiosis duration during the first spermatogenetic wave in 97 C, is strictly consistent with a role of Fignl1 in the regulation of meiosis. The observations reported in mouse osteoblasts, pancreatic cells and especially in *Caenorhabditis* gonads contribute to propose *Fignl1* as an important novel agent of control of meiosis dynamics. Since Fignl1 belongs to the AAA protein family whose function depends on ATP consumption, it is possible that proteins involved in energy supply to other proteins of musculus genomic origin, could malfunction on a spretus factor. This hypothesis does not exclude the degradation kinetics that we suggest as plausible for explaining Fignl1 spretus dysfunction in the musculus context.

In summary, all the data collected in the literature and those obtained by our own experiments converge to present *Fignl1* as a strong candidate for the testicular phenotype of 97 C mice. Indeed, compared to the 11 genes of the refined region, *Fignl1* is the only gene which is expressed in the cell type related to the phenotype, has a known role in a function (especially the passage from mitosis to meiosis in *C. elegans*) related to the phenotype and

Table 1. List of genes and annoted sequences localized between 9.6 and 14.7 Mbp on Mouse Chromosome 11.

Gene symbol	Expression level in B6 testis (arbitrary fluorescence units)	Ratio of expression 97C/B6	Gene name
Vwc2	129	1.75	von Willebrand factor C domain containing 2
Zpbp	7259	0.77	zona pellucida binding protein
LOC48 2554	61	0.49	
Pmv1Rik	10969	0.87	
4930512M02Rik	94	0.36	
Ikzf1 (Zfpn1a1)	137	0.68	IKAROS family zinc finger 1
Fignl1	16618	0.99	Fidgetin-like1
Ddc	75	0.83	Dopa decarboxylase
Grb10	1407	0.76	growth factor receptor-bound protein 10
Cobl	192	1.61	cordon-bleu
LOC68448	134	0.57	

doi:10.1371/journal.pone.0027582.t001

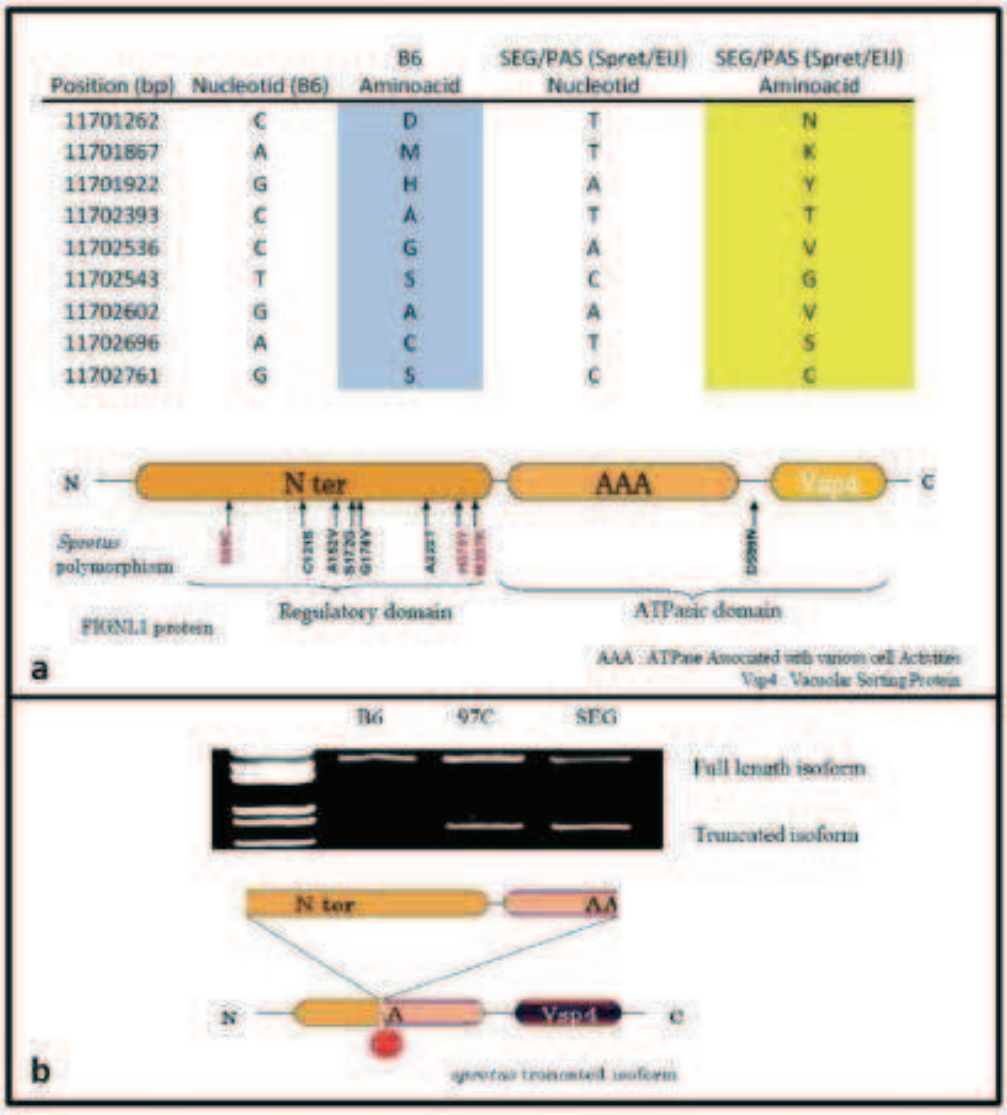


Figure 7. Specific features of spretus Fignl1. a) Polymorphism map of *spretus fignl1* gene. By PCR and sequencing of spretus Fignl1, we could observe 9 coding variants compared to the B6 version. These variants are mainly located (8 of them) in the N terminal region of Fignl1, reported as less conserved. In red are represented variants that could give rise to strong and significant variations in the protein structure or to post-translational modifications. S99C may generate disulphide bonds with other cystein residues, H379Y generates a phosphorylable tyrosine, and M397K is specifically encountered in Mus spretus, by contrast with all other species, including monotremes, birds, amphibians and fish (Figure S3). b) RT-PCR amplification of fignl1 ORF on testis cDNA of B6, SEG and 97C males, with primers spanning the complete ORF. The specific amplification product observed when a spretus allele is present corresponds to a mRNA smaller of about 1400 bp which may generate a truncated isoform of the protein; It is interesting to notice, that this isoform should not be problematic in spretus since it is normally present in this species.
doi:10.1371/journal.pone.0027582.g007

harbors several non synonymous SNPs differentiating the *spretus* alleles from the *musculus* alleles of *Fignl1*. The amount of Fignl1 could be important for regulating meiosis length, this amount may be affected either by an increased synthesis or by a protection against degradation. We did not observe a specific increase in the level of gene expression in 97 C testes; this, in accordance with C. elegans data [19] could lead to speculate that this degradation could be the crucial mechanism. As a hypothesis, it is conceivable for instance that ubiquitinylation systems of Mus musculus are less efficient to degrade Mus spretus versions of Fignl1, leading to an accumulation of the protein able to slow down the meiosis process. This slowing meiosis will result in apoptosis or in an increased occurrence of sperm anomalies, thereby explaining the terato-azoospermia observed in 97 C.

4) How the *spretus* fidgetin like1 allele could affect male meiosis in 97 C mice?

Since the expression level of *Fignl1* is the same in adult 97 C and B6 testis (see Table 1), we ruled out the hypothesis that its promoter/regulation region, from *spretus* origin, does not fit with

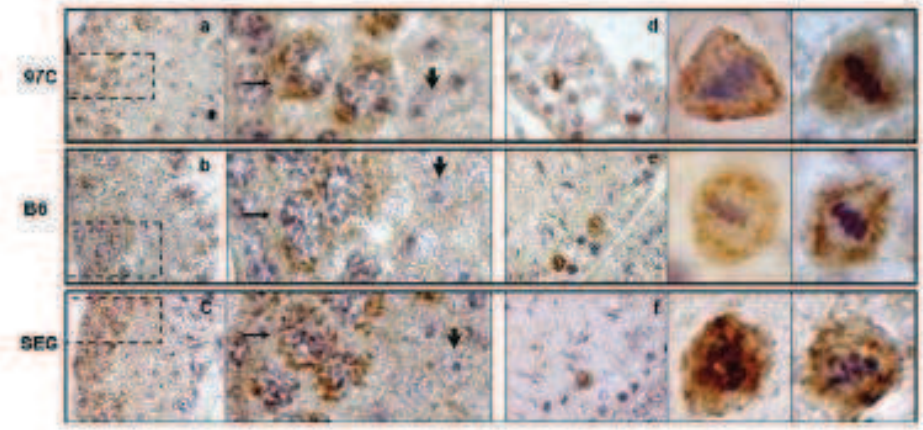


Figure 8. Immunohistochemistry using an anti-FIGNL1 antibody on mouse testes sections. From a cell type point of view, the location of the protein does not appear different between 97 C, B6 and SEG. (a and d) correspond to images from 97 C testes, (b and e) images from B6 testes and (c and f) images from SEG testes taken at objective ×100. (a, b and c) show stages VI-VIII of the spermatogenesis, close to spermiation. The Fignl1 is concentrated in the cytoplasm of mid-pachytene spermatocytes in a crescent-like pattern (see thin arrows in the enlarged views of the dashed area in a, b and c). To note the absence of staining in the spermatids (large arrows in the enlarged views). (d, e and f) show a stage XII with immunoreactive spermatocytes in metaphase. Notice in the enlargements presented at the right side of d, e and f pictures, the very particular location of Fignl1 around the achromatic spindle. This observation is representative of metaphases in B6 mice. In 97 C, a diffuse location of Fignl1 around the chromosomes could also be observed.
doi:10.1371/journal.pone.0027582.g008

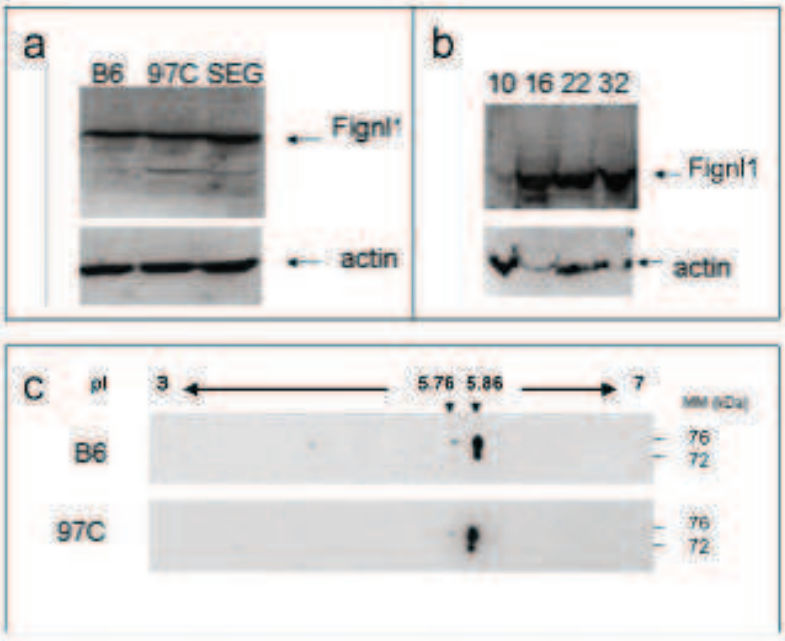


Figure 9. Biochemical characterization of Fignl1 from 97 C, B6 and SEG testes. (a) Western blot revealed by anti-FIGNL1 antibody on extracts of B6, 97 C and SEG testes (50 µg of protein extract per lane). β-actin is taken as an internal control of loading. The Fignl1 band migrates around 74 kDa. There is no obvious difference in size and quantity between the two mouse species, as well as in the 97 C IRCS. The alternate splicing (Figure 7) should generate a ~6 kDa peptide that was not visible in our electrophoresis conditions. (b) Western blot revelation of the Fignl1 during the first wave of the spermatogenesis in B6 testis from 10 to 32 days post partum (DPP). Fignl1 is not detectable at 10 DPP, date of the meiosis entry but is observed at 16 DPP, and afterward, when spermatocytes appear in the tubules. β-actin is taken as an internal control of loading. To note the 7 kD isoform described in the text could not be detected in this 12% polyacrylamide gel. (c) Two-dimensional Western blot of protein extracts of B6 and 97 C testes revealed by anti-FIGNL1. Fignl1 migrates as two principal spots at 76 kDa and 72 kDa for a same 5.86 pI. The 2D migration profile was the same for B6 and 97 C.
doi:10.1371/journal.pone.0027582.g009

transcription factor of B6 origin, that should act on this promoter. Rather, the problem takes its origin from a Fignl1 ^{spretus} dysfunction in the B6 context, caused by evolutionary divergences in the sequence. We found nine non-synonymous polymorphisms between *Fignl1*^{spr} and *Fignl1* ^{spretus} alleles. Eight out of the 9 mutations were found in the N terminal sequence, affecting neither the AAA nor the Vsp4 domains of Fignl1 (figure 7). In a review of 2007, White and Lauring put forward the divergent N-terminal domain utilized by these AAA proteins to interact with adaptor partner proteins [25]. So, it is possible that Fignl1 and its protein partners have co-evolved differently over the 2MYr separating *Mus spretus* and *Mus musculus* species, preventing an efficient interaction between Fignl1 ^{spretus} allelic form and its *Mus musculus* protein partners. To note, a sequence alignment of *spretus* Fignl1 with the sequence of other vertebrate species showed that among the nine amino acid substitutions, H379Y and M397K, touched conserved sequences and was specific to *spretus* Fignl1 (Figure S3).

Another difference between *spretus* and *musculus* Fignl1 was the expression by 97 C and SEG mouse strain testes of a truncated *Fignl1* isoform in addition to the full length isoform which is present in B6 (figure 7b). This truncated isoform corresponds to an alternative splicing between acceptor and donor AG site (position 401 bp and 1706 bp one NM_021891.3 transcript) removing approximately 1400 bp of the coding region. Interestingly, these cryptic splicing sites do exist in the B6 isoform, but seem to be discarded by the splicing machinery, which suggests that some of the synonymous polymorphisms present in the *spretus* isoform either affect or create an exonic splicing silencer element. This truncated isoform might encode a short peptide corresponding to the 60 first amino acids of Fignl1. An *in silico* analysis of this short polypeptide on ELM (Eukaryotic Linear Motif resource) functional prediction site (http://elm.eu.org/) showed that this peptide contains several putative functional signals, and notably a destruction motif, targeted by the anaphase-promoting ubiquitin ligase complex APC/C. The targets of the APC/C are degraded to ensure the correct progression in the cell cycle through mitosis and meiosis [26]. This short polypeptide may act as a competitor of the full length protein in this degradation process. Interestingly, it has been shown that, in *C. elegans*, Fig1 is targeted specifically by the CUL-3/MEL-26/E3 ligase complex resulting in its degradation during meiotic stages [22]. The authors suggest that this degradation might be necessary to ensure the mitotic to meiotic passage of the worm germ cells. Thus a defect in the degradation of Fignl1 during meiosis could lead to a lengthening of the process. Interestingly, this short isoform, when fused with eGFP and transfected in cultured cells, displays a cytoplasmic expression, in a dotted pattern, different from that of eGFP alone, arguing in favor of existence of degradation signal in the N terminus domain of Fignl1 protein (data not shown).

In this study, we have carried out the positional cloning of a gene able to modulate testis weight by interfering with meiosis. For achieving this, we used a mouse model where *Mus musculus* genome is locally substituted by *Mus spretus* genome. The identification of Fignl1 as responsible for the phenotype has been obtained by narrowing down the minimal interval containing the gene on chromosome 11, by eliminating the other genes present in the interval, according to experimental data and bioinformatics and by establishing a parallel with the phenotype described in knocking down the homologous gene (*Fig1*) in *Caenorhabditis elegans*. It is clear that the generation of null mice or even better, a knock-in experiment substituting the spretus *Fignl1* in a B6 genomic context, would directly prove the involvement of the gene. However, our study clearly suggests that Fignl1 has an important role in controlling male meiosis in mice. Dissection of molecular events leading to meiotic impairment in our original model would contribute clarifying this complex spermatocyte prophase I regulation. This model should make it possible to find the molecular bases of the epistatic breakdown that we observe (i.e. meiotic impairment in 97 C but not in SEG or B6 strain), and to discover *Fignl1* genetically interacting genes. Eventually *Fignl1* deregulation and mutation should be investigated in infertile human male presenting with idiopathic meiotic arrest.

Methods

Ethics statement

The experimental procedures were conducted in accordance with the policies of the University and the Pasteur Institute (Paris) and the Guidelines for Biomedical Research Involving Animals. The animals were kept under standard conditions according to the recommendations for the use of animals in experimental designs and according to the "3R" rules. The work was performed under the agreement N/Ref RL-00014323000103O, authorization 75–1463 obtained from the "Direction Départementale des services vétérinaires de Paris".

Animal housing

Parental strains, C57BL/6J (B6) and SEG/Pas, and 97 C1RCS were provided after weaning by The Pasteur Institute (Paris). 1RCS were constructed according to a scheme previously reported [8,18]. Animals were housed at the animal facility of the Cochin Institute, under normal conditions of light/dark cycle, temperature and free access to mouse food and water.

Microsatellite genotyping

DNA was extracted from mouse tail fragments according to a classical procedure [27]. Three microsatellites on MMU6 (D6MIT224, D6MIT321, D6MIT313) and 12 microsatellites on MMU11 (D11MIT304, D11MIT72, D11MIT74 D11MIT129, D11MIT62, D11MIT259, D11MIT204, D11MIT150, D11 MIT133, D11MIT148, D11MIT63 and D11MIT162) were used in order to precise the boundaries of the *spretus* segment in 97 C mouse strain genome and to genotype congenic sub-strains. Primer microsatellites were retrieved from the Mouse Genetic Informatic website of the Jackson Laboratory. PCR was performed using *Taq* DNA Polymerase (New England Biolabs). PCR products were loaded in a 2% nusieve, 5% agarose gel (Cambrex Bio Science Rockland, Inc).

Generation of sub-congenic mice

97 C mice were crossed with B6 individuals in order to obtain heterozygous mice for the *spretus* fragments. A first F2 population was established and genotyped in order to separate the two 97 C *spretus* fragments located on MMU6 and MMU11. Individuals of both sexes harboring only one of these genomic fragments were selected and crossed in order to establish two congenic sub-strains, 97Crc6 and 97Crc11 respectively. Concerning 97Crc11 sub-strains, a second F2 population was established from a cross between 97Crc11 and B6 mouse then genotyped in order to find recombinations inside the spretus fragment. Such recombinant individuals were selected and crossed with B6 mice in order to obtain recombinants of both sexes which were crossed between them to get homozygous recombinant sub-strains.

Whole Testis RNA extraction and cDNA synthesis by reverse-transcription

Total testis RNA from several SEG/Pas and 97 C males was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA)

according to the manufacturer's instructions. RNA extractions from the two testes of six males of each strain were pooled before DNase I treatment (Invitrogen, Carlsbad, CA, USA).

Testis RNA was reverse transcribed to obtain cDNA using either M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) or Roche *Transcriptor* Reverse Transcriptase following manufacturer's protocols.

Mus spretus allelic variant ORFs sequencing

PCR primers (primer sequences available upon request from the authors) were designed, based upon the B6 sequences in 5' and 3' UTR sequences surrounding the ORFs of *Gtf3*, *Zfp9*, *Pus1Rik* and *Rgsl1* genes. These primers were then used to amplify *Mus spretus* gene versions from testis cDNA extracted from SEG/Pas and 97 C males. PCR products were purified and sequenced. *Mus spretus* sequences were BLASTed against the B6 sequences, taken as the reference.

Morphological and histological analysis of the reproductive organs

The male adult mice (8–9 week old) were killed by cervical dislocation and reproductive organs (testes, epididymis and seminal vesicle) were dissected and weighted. For the study of post-natal testis development and first wave of spermatogenesis, juvenile mice were killed by decapitation at day post partum (DPP) 9, 14, 21, 24, 28, 35, 42 and 49 and their testes were recovered and weighted. The adult and juvenile organs were either fixed for histological analysis or deep frozen at −80°C for further protein analysis.

For histological analysis, organs were immersed in DF2 fixative (35% absolute alcohol, 10% acetic acid, 2% formaldehyde, in distilled water) for 24 hr at room temperature then embedded in paraffin. Organs were sectioned (4 µm) then deparaffinized and stained by hematoxylin for histological tubule examinations.

Immunohistochemistry

Testis sections were deparaffinized, rehydrated and then treated either by 0.2% TX-100 in PBS for 5 minutes or by 3×5 minutes microwave in citrate buffer (0.1 M sodium citrate and 0.1 M citric acid, pH 6.0) just before blockade of peroxidase activity by incubation with 3% H_2O_2 for 10 min. After non specific binding site saturation, sections were incubated with anti FIGNL1 antibody (Novus Biologicals) (1/100 or 1/50 in blocking solution) for one hour at room temperature and then with horseradish peroxidase-conjugated anti-rabbit antibody. The sections were finally treated with diaminobenzidine in the dark, washed and then rapidly counterstained with Hemalun de Mayer (RAL) and mounted in Eukitt medium (Labonord). Negative controls were performed by incubating sections with rabbit IgG (1/100) in place of the primary antibody.

Protein extract and Western blot analysis

Protein from adult or juvenile testes were extracted by homogenization of whole testis in Laemmli lysis buffer with β-mercaptoethanol and protein extract were boiled 10 minutes at 95°C before loading on SDS-PAGE 12% gels.

For 2D separation, testis proteins were solubilized in 9 M urea, 4% CHAPS, 100 mM DTT and 2% IPG buffer at room temperature. Extracted proteins were separated in precast 3–10 IPG strips (GE Healthcare) and then in SDS-PAGE 12% gels for the second dimension.

The proteins were electro-transferred to a polyvinylidene difluoride membrane and then probed with rabbit anti-FIGNL1 antibody (1/1000) and HRP labelled anti rabbit IgG. HRP activity was detected using the ECL kit (GE Healthcare) and exposed to Kodak X-OMAT film (Fisher-Scientific).

TUNEL assay

Sections of testes from adult mice were subjected to TUNEL assay. Fixed sections were deparaffinized, rehydrated and demasked by 20 µg/ml proteinase K (RNA grade, Invitrogen Corporation). Apoptotic cells were revealed using in Situ Cell Death detection kit, AP (Roche Diagnostics) that labels the DNA strand breaks with fluorescein-dUTP. Sections were counter-stained with DAPI before two independent fluorescence microscopy examinations. Two hundred twelve and 161 tubule sections were examined for 97 C and B6 mice, respectively. Fluorescent cells (TUNEL positive) were counted in each tubule section on different fields observed at objective ×20.

Statistical tests

The Student T test and the non parametric Kruskall and Wallis test were used to compare mean values and Chi square test for percentage value. Values were considered significant under the threshold of 0.05 ($p<0.05$).

Supporting Information

Figure S1 Histological sections of epididymal duct at 28 and 35 days post partum in 97 C and B6 mice. At 35 DPP, we observed the presence of spermatozoa in the epididymal duct lumen of B6 mice whereas only abnormal/apoptotic round cells are visible in 97 C epididymis. These cellular elements are normally present in the epididymal duct lumen at 28 DPP consequent to the setting up of the first wave of spermatogenesis.
(PDF)

Figure S2 Expressional data of *Pus1* during the first wave of spermatogenesis. RT-PCR amplification of *Pus1* ORF from testis cDNA of B6 mice at 10, 16, 20, 22, 24 and 30 days post partum. The specific amplification product is undetectable at 10, 16, and 20 DPP when germ cells are only represented by spermatogonies and spermatocytes in the tubules. It became observable from 22 DPP, concomitantly with the apparition of spermatids which begun to differentiate in the tubules.
(PDF)

Figure S3 Phylogenetic conservation of Fignl1 across vertebrate species. In blue are represented amino-acids of the B6 type, while in yellow are represented amino-acids of the spretus type. The amino-acids located at positions 379, 397 and 599 are apparently strictly specific of spretus SEG/Pas. It is interesting to notice that S99C is generally specific of non mammal species (except the monotremata).
(PDF)

Acknowledgments

We thank K. Reynard for providing us cat and dog testes and M. Guéroun for chicken testis.

Author Contributions

Conceived and designed the experiments: DLH XM CS DV. Performed the experiments: DLH MV JA JC BP CS. Analyzed the data: DLH CS DV. Contributed reagents/materials/analysis tools: JC BP XM. Wrote the paper: DLH CS DV.

References

1. Leroy JL, Opsomer G, Van Soom A, Goovaerts IG, Bols PE (2008) Reduced fertility in high-yielding dairy cows: are the oocyte and embryo in danger? Part I. The importance of negative energy balance and elevated ovarian luteinizing hormone pulse frequency in the reduction of oocyte and embryo quality in high-yielding dairy cows. Reprod Domest Anim 43: 612–622.
2. Ozaki O, Kadoya S, Koyama H (2010) Factors affecting reproductive performance in high milk-producing Holstein cows. J Reprod Dev 56 Suppl: S61–65.
3. Voigt PH, Fabes C, Hamerin R, Zimmer J (2008) The AZF proteins. Int J Androl 31: 385–394.
4. McLachlan RI, O'Bryan MK (2010) State of the art for genetic testing of infertile men. J Clin Endocrinol Metab 95: 1013–1024.
5. Matzuk MM, Lamb DJ (2008) The biology of infertility: research advances and clinical challenges. Nat Med 14: 1197–1213.
6. Lalonde P, L'Hote D, Serres C, Vaiman D (2009) Mouse models for identifying genes modulating fertility parameters. Animal 3: 35–71.
7. L'Hote D, Lalonde P, Serres C, Montagutelli X, Vreiz RA, et al. (2010) Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics 177: 2321–2333.
8. Burgio G, Stefanski M, Gunner JL, Arnau MR, Panthier JJ, et al. (2007) Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics 177: 2321–2333.
9. Newman TL, Tuzun E, Morrison VA, Hayden KE, Ventura M, et al. (2005) A genome-wide survey of structural variation between human and chimpanzee. Genome Res 15: 1344–1356.
10. L'Hote D, Serres C, Laissue P, Oulmouden A, Rogel-Gaillard C, et al. (2003) Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. Genetics 176: 1909–1921.
11. Mu X, Lee YF, Liu NG, Chen YT, Kim E, et al. (2004) Targeted inactivation of testicular nuclear orphan receptor 4 delays and disrupts late meiotic prophase and subsequent meiotic divisions of spermatogenesis. Mol Cell Biol 24: 5887–5899.
12. L'Hote D, Serres C, Vreiz RA, Montagutelli X, Oulmouden A, et al. (2006) Gene expression regulation in the context of mouse interspecific meiotic genomics. Genome Biol 7: R135.
13. Liu YN, Roy A, Yan W, Burns KH, Matzuk MM (2007) Loss of zona pellucida binding proteins in the acrosomal matrix disrupts acrosome biogenesis and sperm morphogenesis. Mol Cell Biol 27: 6794–6805.
14. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, et al. (2008) Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet.
15. Georgopoulos K, Bigby M, Wang JH, Molnar A, Wu P, et al. (1994) The Ikaros gene is required for the development of all lymphoid lineages. Cell 79: 143–156.
16. Molnar A, Wu P, Largespada DA, Verlaamp A, Scherer S, et al. (1996) The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse. J Immunol 156: 585–592.
17. Erdmann B, Wirbel FF, Florian A, Ryten J, Beyer A, et al. (1991) PAS1, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. Cell 64: 499–510.
18. Frickey T, Lupas AN (2004) Phylogenetic analysis of AAA proteins. J Struct Biol 146: 2–10.
19. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, 3rd, Hurov KE, et al. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316: 1160–1166.
20. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, et al. (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci Signal 3: ra3.
21. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325: 834–840.
22. Lukas-Glaser S, Pintard L, Tyers M, Peter M (2007) The AAA-ATPase FIGL-1 controls mitotic progression, and its levels are regulated by the CUL-3/MEL-26 E3 ligase in the C. elegans germ line. J Cell Sci 120: 3170–3180.
23. Park SJ, Kim SJ, Rhee Y, Byun JH, Kim SH, et al. (2007) Filamin-like 1 gene inhibited by basic fibroblast growth factor regulates the proliferation and differentiation of osteoblasts. J Bone Miner Res 22: 686–696.
24. Schraenen A, de Faudeur G, Thorrez L, Lemaire K, Van Wichelen G, et al. (2010) mRNA expression analysis of cell cycle genes in islets of pregnant mice. Diabetologia 53: 2579–2588.
25. White SR, Lauring B (2007) AAA+ ATPases: achieving diversity of function with conserved machinery. Traffic 8: 1657–1667.
26. Peters JA, Orr-Weaver TL (2008) Regulation of APC/C activators in mitosis and meiosis. Annu Rev Cell Dev Biol 24: 475–499.
27. Baker VV, Shingleton HM, Hatch KD, Miller DM (1988) Selective inhibition of c-myc expression by the ribonucleic acid synthesis inhibitor methotrexate. Am J Obstet Gynecol 158: 762–767.

Article 2

Refined mapping of a Quantitative Trait Locus on chromosome 1 responsible for mouse embryonic death.

(Vatin et al., 2012)

Résumé

Les avortements spontanés à répétition (RSA) sont définis comme au moins trois échecs de grossesse consécutifs au cours du premier trimestre de développement embryonnaire intra-utérin. Il s'agit d'une pathologie fréquente de la grossesse qui affecte 1 à 5% des femmes dans la population générale mais dont les causes sont inconnues dans environ 50% des cas.

L'objectif de cette étude est d'identifier de nouveaux gènes candidats potentiellement responsables des RSA. Les lignées IRCS ont été exploitées en utilisant une approche cartographique et une étude d'association génotype/phénotype. La lignée 66HMMU1 a été identifiée comme présentant une région QTL *Led2* (~30Mb) associée à une augmentation de la mort embryonnaire (Laissue et al., 2009).

Des sous-lignées de 66HMMU1 ont donc été créées et j'ai réalisé la cartographie fine du QTL. L'analyse génétique suggère que le phénotype observé est lié à un dysfonctionnement utérin. J'ai ainsi pu proposer une région de 6 Mb dans laquelle 7 gènes candidats (*Ncl*, *Cab39*, *Eif4e2, Trip12, Psmd1, Cops7b and Usp40)* ont été retenus comme potentiellement responsables du phénotype observé. Ces gènes codent pour des protéines appartenant à différentes voies de signalisation telles que la voie mTOR ou celle de la

dégradation des protéines, dépendante de l'ubiquitine/protéasome et sont impliqués dans plusieurs processus physiologiques tels que l'angiogenèse.

OPEN ACCESS Freely available online

PLOS ONE

Refined Mapping of a Quantitative Trait Locus on Chromosome 1 Responsible for Mouse Embryonic Death

Magalie Vatin[1], Gaetan Burgio[2,3], Gilles Renault[1], Paul Laissue[4], Virginie Firlej[1], Françoise Mondon[1], Xavier Montagutelli[2], Daniel Vaiman[1], Catherine Serres[1], Ahmed Ziyyat[1]*

1 Université Paris Descartes, Institut Cochin Inserm U1016 CNRS UMR 8104, Paris, France, 2 Institut Pasteur, Unité de Génétique des Mammifères, Paris, France, 3 Department of Genetics, Menzies Research Institute, University of Tasmania, Hobart, Australia, 4 Unidad de Genética, Facultad de Medicina, Universidad Del Rosario, Bogota, Colombia

Abstract

Recurrent spontaneous abortion (RSA) is defined as the loss of three or more consecutive pregnancies during the first trimester of embryonic intrauterine development. This kind of human infertility is frequent among the general population since it affects 1 to 5% of women. In half of the cases the etiology remains unelucidated. In the present study, we used interspecific recombinant congenic mouse strains (IRCS) in the aim to identify genes responsible for embryonic lethality. Applying a cartographic approach using a genotype/phenotype association, we identified a minimal QTL region, of about 6 Mb on chromosome 1, responsible for a high rate of embryonic death (~30%). Genetic analysis suggests that the observed phenotype is linked to uterine dysfunction. Transcriptomic analysis of the uterine tissue revealed a preferential deregulation of genes of this region compared to the rest of the genome. Some genes from the QTL region are associated with VEGF signaling, mTOR signaling and ubiquitine/proteasome-protein degradation pathways. This work may contribute to elucidate the molecular basis of a multifactorial and complex human disorder as RSA.

Citation: Vatin M, Burgio G, Renault G, Laissue P, Firlej V, et al. (2012) Refined Mapping of a Quantitative Trait Locus on Chromosome 1 Responsible for Mouse Embryonic Death. PLoS ONE 7(8): e43356. doi:10.1371/journal.pone.0043356

Editor: Shizufumi Ebihara, Nagoya University, Japan

Received October 19, 2011; **Accepted** July 23, 2012; **Published** August 16, 2012

Funding: MV is a PhD student funded by the French Ministry of Research (Doctoral School Gc2D, Université Paris-Descartes). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ahmed.ziyyat@parisdescartes.fr

Introduction

Embryonic development in mammals begins from the female and male interaction which leads to the oocyte fertilization. After 5 to 6 cell divisions inside the zona pellucida, the blastocyst undergoes its development conducing to the implantation in the uterine tissue. The external cells of the blastocyst develop into the placenta, a pivotal organ which allows immune tolerance, bidirectional foeto-maternal exchanges and crucial synthesis of gestational hormones [1]. All these biological processes are required for the survival of every mammalian species, and logically, they underlie a high level of complexity. Dysfunctions in these processes can lead to infertility. In humans it is a considerable public health problem, affecting up to 15% of couples. Due to the number of factors involved in a successful reproductive process, the mechanistics of infertility are far to being completely understood.

At present, although hundreds of mutant mouse models with reproductive phenotypes have been generated [2] and substantial progress has been made in the identification of genetic causes of human infertility, more than 70% of the cases are still considered as idiopathic [3]. Among these, recurrent spontaneous abortion (RSA) (defined by the occurrence of at least three successive pregnancy losses) affects one to five percent of couples [4]. This pathology can be the result of chromosomal anomalies [5], maternal and fetal structural abnormalities [6,7], thrombophilic disorders [8] and autoimmune disorders such as the antiphospholipid syndrome [9]. However, in fifty percent of the cases the etiology remains unknown [10,11]. Up to now, RSA genetic causes have already been explored with variable degrees of success. For instance, in 2006, Kaare et al. analyzed the entire open reading frame of the *Anucleolus* gene (*AMN*) in patients affected by RSA but no causal mutations could be identified [12]. More recently, the study of Mercier *et al.* described a statistical association between the p.Val617Phe mutation of the Janus kinase 2 protein and RSA [13]. All in all, the intrinsic difficulty to genetically dissect mammalian reproductive phenotypes, in which hundreds of genes interact into subtle regulatory networks, has not permitted to identify etiological molecular factors that could explain a significant proportion of infertility cases.

In recent years, in order to overcome these constraints we created an original mouse model of interspecific recombinant congenic strains (IRCS) which permit to localize chromosomal regions associated with complex phenotypes (Quantitative Trait Loci or QTL) [14]. This model is composed of 53 strains of mice which harbor, on average, 2% of *Mus spretus* SEG/Pas genome fixed at homozygous state on *Mus Musculus* C57Bl6/J (B6) genomic background. Using IRCS animals we have previously shown that 3 QTL of embryonic lethality mapped on a unique *spretus* fragment in 3 strains, 66H-MMU13, 66H-MMU1 and 135E. The first, *Led1* in 66H-MMU13 strain on the MMU13 (~2.6 Mb) comprised between the rs12069734 and D13Mit47 polymorphic genetic markers. The second, *Led2* in 66H-MMU1 was analyzed in the present study and the third, *Led3* located on MMU19 in

135E strain encompassing a unique Spretus fragment of 8 Mb located between D19Mit49 and D19Mit137 markers. The 66H-MMU1 strain, which encompasses a unique *spretus* chromosomal fragment located on MMU1 is affected by high levels of embryonic death (24.6%). This strain encompasses a QTL of embryonic lethality (named *Let2*) spreading on 32 Mb and containing 215 genes (143 annotated and 72 predicted) [15].

Here, we present a thorough genetic dissection of *Let2*. For this purpose, we created 15 substrains from 66H-MMU1 animals, which encompass distinct overlapping *spretus* fragments. Using in vivo high frequency ultrasonography to follow the embryonic development, we used an approach of type "phenotype/genotype association" to refine this QTL of embryonic death. We identified, into the *Let2* QTL, one region of approximately 6 Mb, *Let2subA*, which has a main effect on the rate of embryonic death. In addition, we pointed out a second region, *Let2subB*, which could also have a small effect on the phenotype, although statistically not demonstrated.

Materials and Methods

Ethics Statement

Procedures for handling and experimentation were conducted in accordance with the policies of the Paris Descartes University, the Cochin Institute and the Guidelines for Biomedical Research Involving Animals. The experiments were approved by the departmental veterinary services of Paris (approval number: A75 14-02).

Animals

The 66H-MMU1 strain was created at the Pasteur Institute (Paris) by successive crosses of the two parental species *Mus musculus* (C57BL6/J) and *Mus spretus* SEG/Pas (originating initially from Spain). The design of these crosses was reported in a previous work [16]. For this study, 15 new recombinant substrains were generated by backcrosses of 66H-MMU1 with C57BL6/J mice. After weaning, 4 weeks aged mice were maintained in an animal facility of the Cochin Institute (Paris) at normal temperature (21–23°C) and 14 h light/10 h dark photoperiods with free access to water and food.

Microsatellite Genotyping

DNA was extracted from mouse tail fragments by a standard procedure. Eight new microsatellites located on MMU1 (D1Mit439, D1Mit183, D1Mit44, D1Mit383, D1Mit8, D1Mit384, D1Mit255, D1Mit438) were genotyped in order to precise the boundaries of the *spretus* segment present in the 66H-MMU1 genome. Primer microsatellites were retrieved from the Mouse Genome Informatic website (MGI) website of the Jackson Laboratory (www.informatics.jax.org). PCRs were performed using Taq DNA Polymerase (New England Biolabs). PCR products were loaded in a 2% Nusieve, 2% agarose gel (Cambrex Bio Science Rockland, Inc).

Phenotyping: Ultrasonographic Examinations

The gestation was obtained by crossing each IRCS female with a C57/BL6 male. Each female was used one time to collect phenotypic data from the primo-gestation. For each group one female per gestation and the number of animals studied is always >4. C57/BL6 females from the control group were crossed with C57/BL6 males. Substrain's phenotyping was carried out at the small animal imaging facility of the Cochin Institute using high frequency ultrasonography (VEVO 770, Visualsonics, Toronto, Canada). Eight to 12 weeks IRCS females were mated with C57BL/6J males, during a period of 2.5 days. Then, female mice were anesthetized with 1.5% of isoflurane in order to achieve ultrasound examination (Minerve Veterinary Equipment, France). Briefly, a chemical hair remover was used to eliminate abdominal hair. Ultrasonographic contact gel was used to ensure contact between the skin surface and the transducer. Body temperature, electrocardiographic and respiratory profiles were monitored using ultrasound device's integrated heating pad and monitoring device (THM150, Indus Instruments, Webster, TX, USA). Examinations were performed using 2 different high frequency probes depending on the size of the embryos: a 60 MHz transducer for early stages of development (RMV708) and a 40 MHz (RMV704) transducer for late developmental stages.

In order to follow the gestation in vivo, three ultrasonographic examinations were performed at three time points (between E7 and E16). During each examination, we assessed the number of implanted embryos in each uterine horn as well as their status (alive or dead) was assessed. For each gestation, the embryonic lethality rate was calculated as the number of resorbed embryos in both horns reported to the total number of implanted embryos.

RNA Extraction and Gene Expression Arrays

Females from IRCS and B6 strains were crossed with B6 males. Each pregnant female was subjected to ultrasonographic examinations in order to precisely determine the embryonic developmental stage. Female mice were euthanized by cervical dislocation and tissues were taken at E12.5. Total RNA of uterine tissue from six mice of the IRC substrain of interest was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. Similarly, six B6 were used to extract total RNA. In order to duplicate microarray hybridizations of samples from uterine tissues, two pools of three RNA extractions were created for both IRCS and B6 animals. One microgram of RNA from IRC substrains of interest and the B6 controls was used for hybridization on a NimbleGen expression array. cDNA synthesis, DNA end-labeling, hybridization, scanning, and data normalization were performed at the genomic/transcriptomic platform of the Cochin institute. The average fluorescence values for each transcript were collected chromosome per chromosome for each analyzed strain. To evaluate gene expression modifications in IRC strains, these fluorescence levels (considered as expression values) were divided gene per gene by the corresponding ones from B6 which were taken as a reference. The results of the gene expression were deposited at GEO (NCBI) with the accession reference GSE32460.

cDNA Synthesis by Reverse-transcription and Quantitative RT-PCR

After RNA preparation, the total RNA was treated with DNase I (Invitrogen Life Technologies) for 10 min at room temperature followed by inactivation with EDTA (Sigma-Aldrich). Total uterus RNA was reverse transcribed to obtain cDNA using M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) following manufacturer's protocols.

Quantitative PCR was carried out using fast SYBR Green Master Mix (Applied Biosystems) and a real time PCR system (Light Cycler 1.5, Roche Diagnostics, Division Applied Sciences, Meylan, France) according to standard PCR conditions. To validate the primers used in qRT-PCR, four pairs of primers were tested for each gene and four housekeeping genes were also tested to choose the reference gene (Table S1 in supplemental data). For quantitative calculations, values were normalized to mouse *Gapdh/hr A* expression. Primer sequences are listed in Table S1.

Statistical Analysis

Results are expressed as mean±SEM calculated from the variation between individual female. The statistical significance of the differences observed between the mean values of IRCS and the control group (C57BL/6J) samples was evaluated by t-test using the Bonferroni-corrected levels. As we used 7 substrains, a p value <0.007 (0.05/7) was considered as significant. Statistically significant results are labeled as follows in all figures: *: p<0.05; **: p<0.01; ***: p<0.001.

Results

Creation of 66H-MMU1 Substrains

We started our study using the 66H-MMU1 strain which harbors a high rate of embryonic lethality (24.6%) caused by the *Led2* QTL [15]. This QTL of ~32 Mb, which was initially delimited by D1Mit50 (87.0 Mb) and rs6259837 (119.1 Mb) markers located on chromosome 1, corresponds to a *spretus* fragment carried by the 66H-MMU1 substrain. However, a uncertainty of ~6.4 Mb existed at the proximal boundary of this QTL since the interval comprised between D1Mit134 (80.6 Mb) and D1Mit50 (87.0 Mb) markers corresponds to this distance and the breakpoint is located somewhere between these two markers. Indeed, D1Mit134 and D1Mit50 allele markers are of B6 and *spretus* natures respectively (http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm). In an attempt to precise the position of the breakpoint, we genotyped 8 novel markers located on this region that permitted to reduce the recombination region to a 2.5 Mb interval comprised between 84.5 Mb and 87 Mb (markers D1Mit430 and D1Mit50 respectively, see Figure 1). Then, we initiated a fine mapping approach using fifteen recombinant substrain issued from 66H-MMU1 animals. In each of these strains, a crossing-over fragmented the original DNA region of *spretus* origin that was initially present in the 66H-MMU1 strain (Figure 1). Among 15 starting strains, seven survived and were available for our study (recombinants R3, R4, R5, R6, R10, R13 and R14).

In vivo Phenotyping

Females of the different recombinant substrains were crossed with B6 males and their gestation was followed up using *in vivo* ultrasonography as previously described [15]. Animals obtained from these crosses have a placenta which is heterozygous for all the genes located on the fragment of *spretus* origin (B6/SEG), while the uterus was homozygous SEG/SEG for the same *spretus* fragment. The control group was obtained by crossing males and females of the B6 strain. A total of 97 gestations (31 and 66 of B6 and IRCS types respectively) were analyzed. For each gestation, we counted the number of implanted and resorbed embryos during three ultrasonographic examinations. There was no correlation between embryonic death and the position inside the womb, which suggests that the death of one embryo did not have deleterious repercussions on the contiguously implanted structures (Figure 2).

We noted a strong variability in the percentage of embryonic death between the different substrains (Table 1), while apparently the number of implanted embryo was not significantly different. The strains R4, R6, R10, R13 and R14 (group 1) presented a percentage of embryonic death that was not statistically different from that of B6 control animals (10±4%, 19±5%, 8±3%, 12±5% and 16±8% for the five strains, respectively, versus 12±2% in the B6 parent). Note that within this group, R6 is distinguished by the highest rate of embryonic death. By contrast, R3 and R5 strains (group 2) presented a percentage of embryonic death (27±5% and 29±6% respectively) significantly higher than B6 control at $p = 0.0013$ and $p = 0.0045$ respectively (Table 1). Since the mean number of implanted embryos was not different between the substrains compared to the B6 control (Table 1), we deduced that the increase in embryonic death observed for R3 and R5 IRCS was caused by post-implantation events.

QTL Fine Mapping

In order to refine *Led2* localization, we realized an analysis by genotype/phenotype segregation. R3 and R5 strains (which exhibit the phenotype) shared a large *spretus* region (>84.5 Mb to 90.5 Mb) with the R6 strain (which does not display the embryonic resorption phenotype) and the rest (until <100.3 Mb) is also shared with the other strains (R4, R10, R13 and R14) which are not affected. This configuration suggests that two *spretus* regions, shared by R3 and R5 strains and not present together in the other strains, seem to be necessary to explain the apparition of the phenotype in R3 and R5. We defined a first *spretus* subfragment called *Led2min4* which encompasses D1Mit50 to

Markers	D1 Mit 438	D1Mit 50	D1Mit 81	D1Mit 415	D1Mit 305	D1Mit 440	D1Mit 488	rs622 012	D1Mit 10	rs369 2309	rs134 76085	rs368 5663	gnf01 -104-109	rs366 3003	rs359 7152	rs625 9837	rs134 76083
Position	84.5 Mb	87.0 Mb	87.6 Mb	88.2 Mb	90.5 Mb	90.6 Mb	91.9 Mb	91.9 Mb	92.5 Mb	95.1 Mb	100.3 Mb	102.3 Mb	106.1 Mb	110.6 Mb	115.5 Mb	119.1 Mb	121.2 Mb
R10	B	B	B	B	B	B	B	B	B	S	S	S	S	S	S	S	B
R14	B	B	B	B	B	B	B	B	S	S	S	S	S	S	S	S	B
R13	B	B	B	B	B	B	S	S	S	S	S	S	S	S	S	S	B
R4	B	B	B	B	B	S	S	S	S	S	S	S	S	S	S	S	B
R6	B	S	S	S	S	S	S	S	S	B	B	B	B	B	B	B	B
R3	B	S	S	S	S	S	S	S	S	S	B	B	B	B	B	B	B
R5	B	S	S	S	S	S	S	S	S	S	S	B	B	B	B	B	B

Figure 1. Genomic structure of the IRCS mice in the region of interest of chromosome 1. The map presents the genomic background, spretus or musculus, of the 7 recombinant substrains (Rc) in the chromosomal region corresponding to *Led2* QTL. Recombinant strains were generated at the Pasteur Institute (Paris) from 66HMMU1 strain by recombination events inside the MMU1 *spretus* segment. These strains have been genotyped using 24 polymorph markers. Marker positions are given in megabase pairs (Mb). "S" corresponds to the marker in a spretus homozygous form and "B" to the marker in a musculus (B6) homozygous form. The two minimal spretus regions (*Led2min4* with main effect and *Led2min8* with probable weak effect) responsible for the phenotype of interest are highlighted in gray and in gray hatched when coexisting in the same substrain.
doi:10.1371/journal.pone.0043356.g001

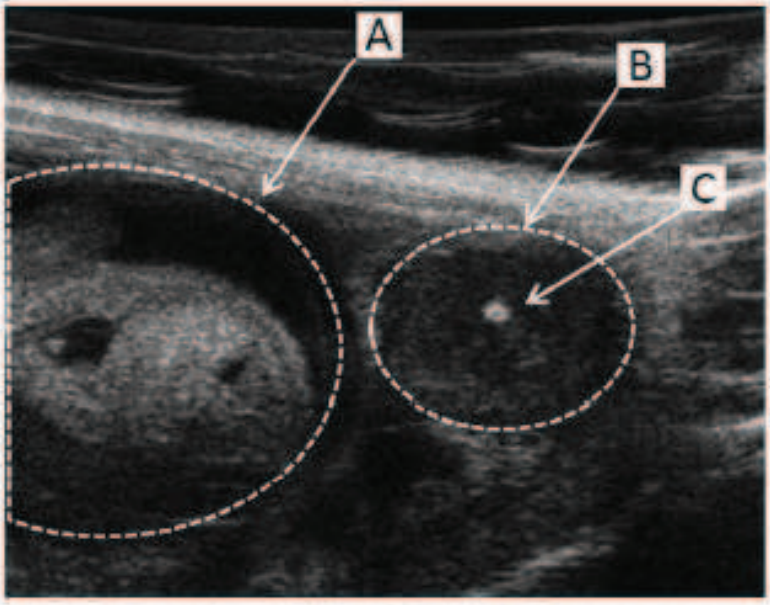


Figure 2. Ultrasound biomicroscopic *in vivo* observation of the embryonic development. During the gestation of B6 and IRC mice the embryonic development was followed up by an *in vivo* ultrasonic method. The viability of developing embryo (A) was assessed by the presence of heartbeats and a positive umbilical cord Doppler. Dead embryos (B) displayed a central highly echogenic zone (C) corresponding to the embryonic resorption.
doi:10.1371/journal.pone.0043356.g002

D1Mit305 region (>84.5 Mb to <90.5 Mb) and a second region called *Lel2wisB* located at the rs3692309 marker (>92.5 Mb to <100.3 Mb) (see gray boxes in Figure 1). When these two *spretus* regions (*Lel2wisA*) are separated as in R6 (that contains *Lel2wisA* only) or in R4, R10, R13 or R14 (*Lel2wisB* only) the phenotype of embryonic death is absent. The presence of the two *spretus* regions seems indispensable to permit the manifestation of the phenotype, it's the case for R3 and R5 (gray hatched boxes in Figure 1). To statistically prove the presence of these two QTLs, *Lel2wisA* and *Lel2wisB* each one responsible for a part of the effect on the phenotype, and an eventual epistatic interaction between these two QTLs able to increase the embryonic death, we compared several recombinants among themselves, R6 bearing *Lel2wisA*, R4 bearing *Lel2wisB* and R3 (or R5) bearing these two *spretus* regions. The results of statistical t-tests are shown in Figure 3. When we compared the mean rate of embryonic death between R4 and R3 (or R5), the statistical result (significant difference at P≤0.01) proved the presence of *Lel2wisA* QTL. By contrast, the comparison between R6 and R3 (or R5), did not statistically indicate the presence of *Lel2wisB* QTL. However the embryonic death rates of R3 and R5 both have a tendency to be higher than that of R6 (Figure 1) suggesting a possible very small effect of *Lel2wisB* on the phenotype. In the same way, the difference in embryonic death rate between R4 (*Lel2wisB* only) and R3 (or R5) was 17%–19% which is comparable to 15%–17% difference between B6 (with no spretus regions) and R3 (or R5) also suggesting a nil or very small effect of *Lel2wisB*. In consequence this result did not support the presence of an epistatic interaction

Table 1. Statistical analysis of the embryonic resorption phenotype.

Strains	Number of gestations	Implanted embryos; Mean ± SEM (p value)	Embryonic resorption rate; Mean ± SEM (p value)
B6	31	7.90±0.41	12%±2
R3	13	7.46±0.50 (p = 0.2717; NS)	**27%±5 (p = 0.0013)**
R4	15	7.20±0.71 (p = 0.1845; NS)	10%±4 (p = 0.3021; NS)
R5	4	9.50±0.50 (p = 0.0918; NS)	**29%±6 (p = 0.0045)**
R6	6	8.50±0.99 (p = 0.2848; NS)	19%±5 (p = 0.1158; NS)
R10	8	9.12±0.51 (p = 0.0828; NS)	9% ±3 (p = 0.2188; NS)
R13	12	8.83±0.45 (p = 0.1045; NS)	12%±5 (p = 0.4349; NS)
R14	8	8.75±0.70 (p = 0.1759; NS)	16%±8 (p = 0.2351; NS)

Comparison between IRCs and B6 control using t-test with Bonferroni-corrected level.
(NS: non-significant).
doi:10.1371/journal.pone.0043356.t001

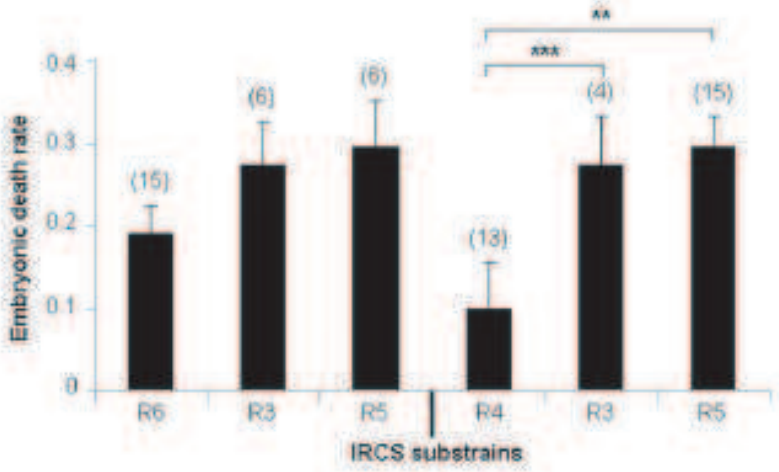


Figure 3. Statistical comparisons between embryonic death rates of IRCS. The mean of embryonic death rate (±SEM) for (n) gestations is presented for four substrains containing different regions of spretus origin: R6 (containing *Led2minA* only), R3 and R5 (containing Led2minA and Led2minB), and R4 (Led2minB only).
doi:10.1371/journal.pone.0043356.g003

between *Led2minA* and *Led2minB* regions, but a *Led2minB* additive effect could be revealed by increasing sample size of this "QTL" representative strains. For this raison, the genes present in these two regions, *Led2minA* and *Led2minB* are listed in Table 2.

Is Placenta or Uterus Responsible for the Led2min Effect?

Since R3 and R5 mice did not harbor any obvious developmental anomaly or pathology, excepting for some embryonic death, it was reasonable to suspect that placental and/or uterine dysfunctions could be responsible for the embryonic lethality increase. Thus, we initiated a genetic approach in order to identify in which of these two organs dysfunction could be related with the phenotype. In the previous set of experimental crosses, IRCS females were mated with B6 males. Genetically, this permitted the co-existence of heterozygous foeto-placental alleles (B6/SEG) and homozygous uterine alleles (SEG/SEG) within the same genomic region (*spretus* fragment). Conversely, we performed reverse crosses (♀B6 × ♂IRCS), giving a heterozygous placenta for the genes of the fragment, but a B6 homozygous uterus. In this situation, the would-be disorders ought to find their origin exclusively from a placental-fetal/embryonic defect, caused by the *spretus* state of the MMU1 fragment, but not from a B6 womb defect. In this optic we realized the cross ♀B6 with ♂R3 (IRCS group2). We observed that the mean of embryonic resorption rate (±SEM) was 0.07±0.04 and not significantly different from the control (♀B6 × ♂B6: 0.12±0.02; p = 0.118) whereas the inverse crossing, leading also to a heterozygous foeto-placental complex implanted in homozygous spretus uterus (R3), produced a significantly higher embryonic resorption rate (0.27±0.05, p = 0.001) (Figure 4). From this last observation, we concluded that a uterine dysfunction is very likely at the basis of the observed phenotype.

Assessing of Differentially Expressed Genes in Uterus

We analyzed the expression level of uterine genes in pregnant R3 females compared to those from B6 control animals at E12.5, an important time point when most resorption occurred during our study. For this purpose, we hybridized cDNA synthesized from RNA uterine tissue to Nimblegen mouse microarrays. The Nimblegen arrays interrogated a total of 25,631 mouse transcripts. Gene expression levels were quantified by fluorescence intensity assessing and ranged from 20 to >50,000 arbitrary units of fluorescence (AUF) (mean value ~5,250 AUF). These results were highly reproducible since they showed strong correlations between experimental duplicates (r = 0.967 for B6 and 0.963 for R3). Thus, for subsequent analysis, we took the average of both values for each transcript. We first focused on transcripts with fluorescence levels higher than 100 AUF, we assumed that values under this threshold were very close to background signals and, with this threshold, we selected 10,005 transcripts (70.6% of the total). We considered a gene as differentially expressed if a two-fold difference of expression (up or down) was observed. Consequently, 3,436 (19% of the expressed uterine genes) transcripts were modified in R3 uterus when compared to those expressed in B6 (Table 3). A similar number of repressed and induced genes was observed (10.9% and 8.1%, respectively). This deregulation was identified over all the genome. However a significantly higher proportion of genes were deregulated when only the MMU1 *spretus* fragment or the two *Led2min* regions were considered (Table 3).

In order to validate the differential gene expression obtained by the microarray analysis, we checked 7 genes of the *Led2minA* QTL region by quantitative RT-PCR. As shown in Table 4 we obtained a very good agreement between microarray and qPCR results.

Considering that the uterine dysfunction can take its origin from a deregulation of the gene expression or/and non-synonymous coding polymorphisms accumulated during independent evolutive processes of *Mus musculus* and *Mus spretus* species, we listed the genes of the *Led2min* QTL corresponding to these criteria and thus potentially involved in embryonic resorption. Finally we considered the transcripts with an expression level >500 AUF and either exhibiting a deregulation (R3/B6 expressional ratio >2 or <0.5) and the presence of non synonymous polymorphisms between *Mus musculus* and *Mus spretus* (provided by SANGER database: http://www.sanger.ac.uk/) (Table 5).

Then, we searched whether deregulated transcripts could be grouped into functional clusters using DAVID database [17], considering 1750 transcripts with a threshold of >500 AUF in the expression level. This analysis led to the identification of five functional groups of genes and signaling pathways that were deregulated, such as ribosome protein genes (p value: 0.00054), endocytosis process (p value: 0.0027), VEGF (vascular endothelial

Table 2. Gene list in the minimal *Led2min* region.

Start of the gene	End of the gene	Gene symbol	Gene name
Led2min4			
84500220	84718814	Thpt2	thyroid hormone receptor interactor 12
84718680	84779509	Fbxo36	F-box protein 36
84785720	84814108	Slc16a14	solute carrier family 16 (monocarboxylic acid transporters), member 14
84906580	84918163	LOC665317	
84963431	84986878	A630032D15Rik	
84990856	84995050	A630040F14Rik	
85038928	85138724	LOC624094	
85079502	85092666	LOC665376	
85122702	85140902	LOC620078	
85138618	85186556	LOC665338	
85232406	85345775	LOC665405	
85408049	85429975	Sp110	nuclear body protein
85431873	85476208	LOC434494	
85481184	85532782	Sp100	nuclear antigen Sp100 Gene
85534251	85567800	A630010G21Rik	
85568118	85625284	LOC637657	
85626533	85661536	Gstt2	glutathione S-transferase, theta 2 (rat)
85722735	85739884	Rnf26	ring finger protein 26
85759653	85763330	4933407L21Rik	
85777037	85792178	Gpr55	G protein-coupled receptor 55 Gene
85852563	85855819	Spata3	spermatogenesis associated 3 Gene
85877133	85886226	2810459M11Rik	
85895780	85910466	Psmd1	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1
85930441	85942088	Htr2b	5-hydroxytryptamine (serotonin) receptor 2B
85965962	86010425	Armc9	armadillo repeat containing 9 Gene
86013466	86013476	B3galt7	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7
86154437	86156604	LOC383536	
86175869	86190626	Ncl	nucleolin
86213405	86219312	Nmur1	neuromedin U receptor 1
86253499	86259229	1700019D13Rik	
86357906	86361869	Ptma	prothymosin alpha
86374185	86413672	Pde6d	phosphodiesterase 6D, cGMP-specific, rod, delta
86418270	86437671	Copz1b	COPI (constitutive photomorphogenic) homolog, subunit 7b (Arabidopsis thaliana)
86497461	86501743	Nppc	natriuretic peptide type C
86533013	86581269	4930429A22Rik	Dis3l2 : DIS3 mitotic control homolog (S. cerevisiae)-like 2
86739906	86742882	LOC620213	
86917560	86921982	Alpl6	Alpp2 alkaline phosphatase, placental-like 2
86923174	86935277	Alpi	alkaline phosphatase, intestinal
86956178	86959083	Alpl	alkaline phosphatase 3, intestine
86979825	86986198	Ece1	endothelin converting enzyme-like 1
87014464	87019578	1700027L20Rik	
87021607	87038277	Chrnd	cholinergic receptor, nicotinic, delta polypeptide
87039580	87042806	Chrng	cholinergic receptor, nicotinic, gamma polypeptide
87045084	87071669	Eif4e2	eukaryotic translation initiation factor 4E member 2
87095534	87141962	Efhd1	EF hand domain containing 1
87158229	87262817	Tiam1	Cdyl2 GRB10 interacting GYF protein 2
87301371	87305655	3110079O15Rik	
87308004	87341517	Ngef	

Table 2. Cont.

Start of the gene	End of the gene	Gene symbol	Gene name
89425720	89428908	Neu2	neuraminidase 2
89451549	89455873	Inpp5d	inositol polyphosphate-5-phosphatase D
89587241	89623693	Atg16l1	autophagy-related 16-like 1 (yeast)
89634850	89676328	Sag	retinal S-antigen
89776291	89851972	Usp40	ubiquitin specific peptidase 40
89926245	90050968	Ugt1a7c	UDP glucuronosyltransferase 1 family, polypeptide A7C
89965979	90050174	Ugt1a6a	UDP glucuronosyltransferase 1 family, polypeptide A6A
90031781	90050968	Ugt1a2	UDP glucuronosyltransferase 1 family, polypeptide A2
90035905	90085919	Dnajb5	DnaJ (Hsp40) homolog, subfamily B, member 5
90094279	90108891	6430706D22Rik	
90134452	90223022	Trpm6	transient receptor potential cation channel, subfamily M, member 6
90281189	90257809	Spp2	secreted phosphoprotein 2
90311041	90341257	Ghyl	glutamine repeat protein 1
90432788	90533814	Arl4c	ADP-ribosylation factor-like 4C
Led2mm8			
92905564	92910197	LOC483332	
92997268	92974343	Col6a2	collagen, type VI, alpha 2
92745512	92780165	Mlph	melanophilin
92755518	92764416	LOC623509	
92786540	92804026	Rab17	RAB17, member RAS oncogene family
92889512	92959524	Lrfn1	leucine rich repeat (in FLII) interacting protein 1
92975508	93001201	Gm817	RNA binding motif protein 44
93010446	93054065	Ramp1	receptor (calcitonin) activity modifying protein 1
93080775	93082727	LOC623580	
93080971	93116510	Ube2f	ubiquitin-conjugating enzyme E2F (putative)
93128748	93151480	Scly	selenocysteine lyase
93152480	93176709	Gm656	espin like
93161478	93192810	4931429F02Rik	(Klhl30) kelch-like 30 (Drosophila)
93196665	93204619	BC056501	(Fam132b) family with sequence similarity 132, member B
93205296	93229169	Jkap	integrin-linked kinase-associated serine/threonine phosphatase 2C
93232188	93258577	1700020N15Rik	
93241588	93249828	Hes6	hairy and enhancer of split 6 (Drosophila)
93246582	93269702	Per2	period homolog 2 (Drosophila)
93329082	93358670	Traf3ip1	TRAF3 interacting protein 1
93370970	93390814	Asb1	ankyrin repeat and SOCS box-containing 1
93681682	93678483	Twist2	twist homolog 2 (Drosophila)
93763139	93976299	Hdac4	histone deacetylase 4
94270118	94304164	Ndufa10	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 10
94310086	94311025	Olfr1416	olfactory receptor 1416
94321223	94322159	Olfr1415	olfactory receptor 1415
94341498	94342452	Olfr1414	olfactory receptor 1414
94403578	94404550	Olfr1413	olfactory receptor 1413
94416737	94419053	Olfr1412	olfactory receptor 1412
94426026	94427868	Olfr1411	olfactory receptor 1411
94436244	94439213	Olfr1410	olfactory receptor 1410
94450318	94451515	Olfr12	olfactory receptor 12
94467550	94472801	Myeov2	myeloma overexpressed 2
94474628	94479347	Otos	otospiralin
94662091	94691862	Gpc1	glypican 1

Table 2. Cont.

Start of the gene	End of the gene	Gene symbol	Gene name
94 705 834	94 793 812	*Ankmy1*	ankyrin repeat and MYND domain containing 1
94 757 194	94 759 826	*0710001E24Rik*	(Dusp28) dual specificity phosphatase 28
94 761 810	94 795 866	*Rnpep1l*	aminyl aminopeptidase (aminopeptidase B)-like 1
94 764 815	94 773 854	*Capn10*	calpain 10
94 809 525	94 851 588	*Gpr35*	G protein-coupled receptor 35
94 853 750	94 864 267	*Aqp12*	aquaporin 12
94 864 883	94 901 228	*Kifc1*	kinesin family member 1A
94 906 550	94 916 909	*Agxt*	alanine-glyoxylate aminotransferase
94 961 761	94 991 354	*2510007B09Rik*	
94 990 150	95 061 480	*B930010M06Rik*	
95 062 548	95 131 471	*Sned1*	sushi, nidogen and EGF-like domains 1
95 129 616	95 136 276	*Mterfd2*	MTERF domain containing 2
95 139 842	95 167 980	*Pask*	PAS domain containing serine/threonine kinase
95 174 050	95 194 024	*Ppp1r7*	protein phosphatase 1, regulatory (inhibitor) subunit 7
95 204 801	95 239 906	*Tmem16g*	Ano7: anoctamin 7
95 236 345	95 301 014	*Hdlbp*	high density lipoprotein (HDL) binding protein
95 252 098	95 252 962	*LOC521662*	
95 304 449	95 340 136	*sept1-02*	septin 2
95 342 530	95 452 188	*Fgd2*	FERM, RhoGEF and pleckstrin domain protein 2
95 452 830	95 466 856	*Stk25*	serine/threonine kinase 25 (yeast)
95 519 099	95 524 168	*Bok*	BCL2-related ovarian killer protein
95 533 716	95 565 044	*Thap4*	THAP domain containing 4
95 584 586	95 613 085	*Atg4b*	autophagy-related 4B (yeast)
95 643 900	95 663 251	*Dtymk*	deoxythymidylate kinase
95 653 470	95 665 267	*Ing5*	inhibitor of growth family, member 5
95 665 598	95 682 580	*D2hgdh*	D-2-hydroxyglutarate dehydrogenase
95 691 749	95 705 800	*Gal3st2*	galactose-3-O-sulfotransferase 2
95 707 243	95 816 522	*LOC666 009*	
95 820 514	95 841 547	*LOC519 597*	
95 855 898	95 855 974 0	*Nrx4*	sialidase 4
95 868 710	95 882 959	*Rtkn2*	programmed cell death 1
97 144 149	97 165 471	*2510044 D20Rik*	Fam174a : family with sequence similarity 174, member A
97 418 057	97 469 000	*St8sia4*	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4
98 425 890	98 475 640	*Gm1 633*	predicted gene 1633
98 549 136	98 702 541	*Slco4c1*	solute carrier organic anion transporter family, member 4C1
98 702 544	98 708 307	*9530 060 107*	
98 736 581	98 827 966	*Slco6b1*	solute carrier organic anion transporter family, member 6b1
98 838 854	98 956 709	*Slco6c1*	solute carrier organic anion transporter family, member 6b1
98 969 265	99 133 860	*LOC634 551*	
99 176 834	99 301 409	*LOC634 546*	
99 474 308	99 483 582	*D15Ertd622e*	DNA segment, Chr 1, ERATO Doi 622, expressed
99 536 554	99 591 955	*Pik3cg2l1*	Ppip5k2: diphosphoinositol pentakisphosphate kinase 2
99 609 805	99 623 575	*4930429M06Rik*	Gin1: gypsy retrotransposon integrase 1
99 651 499	99 702 591 9	*Pam*	peptidylglycine alpha-amidating monooxygenase
100 251 529	100 859 787	*Slco6d1*	solute carrier organic anion transporter family, member 6d1

doi:10.1371/journal.pone.0043356.t002

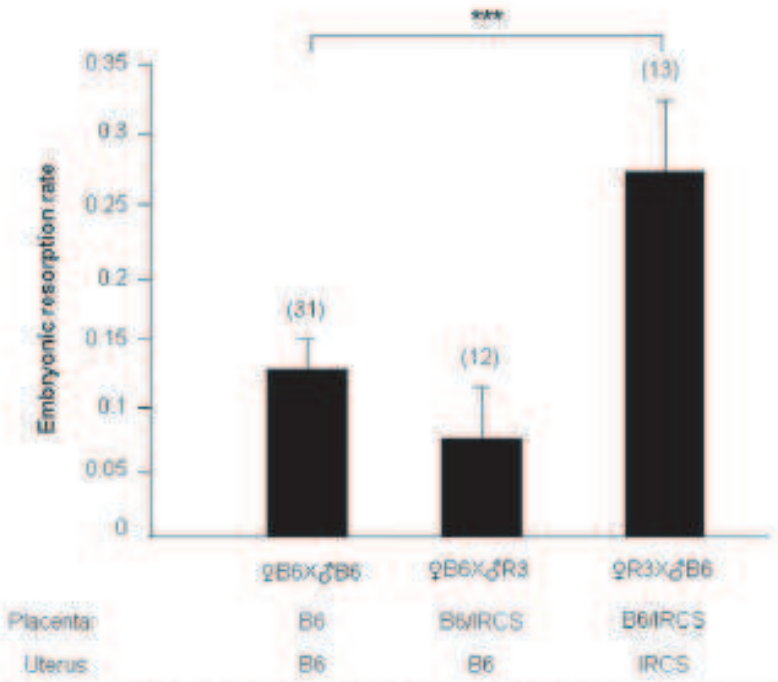


Figure 4. Embryonic resorption rate in function of the type of crossing realized with IRCS and B6 mice. The results of different crosses (♀IRCS × ♂B6, ♀B6 × ♂IRCS and ♀B6 × ♂B6) are presented as the average (±SEM) of the embryonic resorption rate for (n) gestations.
doi:10.1371/journal.pone.0043356.g004

growth factor) signaling (p value: 0.007), chemokine interactions (p value: 0.011) and mTOR (mammalian Target Of Rapamycin) signaling pathway (p value: 0.014).

Discussion

In the reproductive processes, as in others, hundreds of genes interact into subtle regulatory networks, and this complexity does not permit to easily identify the molecular factors of dysfunctions leading to infertility cases. Moreover, when our interest is turned towards the human clinic, the study of factors involved in reproductive defects is particularly challenging due to obvious ethical constraints, which rends obligatory the use of animal models. However although hundreds of mutant mouse models with infertility/hypofertility phenotypes have been generated [2], the genetic causes of infertilities are far from being elucidated in their whole [3]. This is the case for the RSA pathology which affects a non negligible percentage of the population (1 to 5%) and for which the genetic origin(s) is still little documented. In the aim to identify new genes responsible for embryonic lethality, we used a mouse model of interspecific recombinant congenic strains (IRCS). Although during the gestation/pregnancy development, mice and humans do not establish exactly the same system of placentation, similarities strong enough between these two species exist, making the mouse model useful to identify genes involved in humans. Indeed, in several mouse models involving the complement system [18,19], it has been clearly shown that there is a continuum between embryo resorption and placental diseases, since the same mice have these resorption and preeclampsia-like symptoms [20]. It is also known that C3 defects are clearly linked with VEGF defects, thus inducing defective placentation, leading in the most extreme cases to embryonic death. Therefore, mouse models of embryo resorption via known deregulations of complement system are proved to be suitable models of the human continuum placental vascular disease-spontaneous resorption. A well studied mouse model of immunologically mediated peri-implantation pregnancy loss that shares features with human recurrent miscarriage is derivated from DBA/2-mated CBA/J mice (CBA/J × DBA/2) [21,22,23]. Indeed embryos derived from mating CBA/J females with DBA/2 males showed an increased

Table 3. Number of transcripts modified in uterus of R3 IRCS compared to C57BL6/J.

Genomic region	Expressed Transcripts (>100 AUF)	Deregulated transcripts (2-fold threshold)
Total genome	16,048	3,486 (19%)
Spretus fragment in MMU1	144	53 (36%)
Les2minA (84.5–90.5 Mb)	55	24 (44%)
Les2minB (95.1 Mb to < 100.3 Mb)	25	8 (32%)

doi:10.1371/journal.pone.0043356.t003

Table 4. Microarray validation by RT-QPCR on 7 genes of the QTL region.

Gene symbol	Microarrays		RT-PCR		
	Expression level in B6 uterus (AUF)	Ratio of expression R3/B6	Ratio of expression R3/B6	% PCR efficiency	Rsquare value
Nr1	1555	0.16	0.42349187	101%	0.9986
Cab39	1800	4.5	5.48093265	102%	0.99
Prmd1	34109	0.90	0.85502502	104%	0.9934
Copz1b	12535	0.62	0.66533846	105%	0.9999
Elf4c2	652	0.5	0.50713771	105%	0.9919
Usp40	2154	0.21	0.29715833	100%	0.9992
Tsipt 2	12096	0.77	0.60419540	101%	0.9991
Cyclophilin A (Reference gene)	38018	0.95		99%	0.9994

doi:10.1371/journal.pone.0043356.t004

frequency of resorption (29.4±6.5%), more than three times greater than that seen within these and other strains or strain combinations (CBA/J × CBA/J: 8.9±5.1%; CBA/J × BALB/c: 8.2±5.6%; DBA/2 × DBA/2: 8.5±6.6%; n = 6–32 mice/group; CBA/J × DBA/2 vs. others, $p<0.01$) [18]. Spontaneous resorption in the CBA/J × DBA/2 model is attributed to NK cells, macrophages, and Th1-type cytokines, and represent a rejection of the semiallogeneic feto-placental [24]. Murine resorptions are characterized by focal necrosis at the junction of the fetal trophoblast with decidua, an infiltrate of polymorphonuclear leukocytes (with some lymphocytic cells) at sites of necrosis and along the walls of large vessels in decidua, and by thrombosis and hemorrhage [25,26,27,28]. Infiltration begins on day 6.5 of gestation, 2 days after implantation has occurred, and abortions begin after day 8.5 of pregnancy [24,28].

In the present work we used a mouse model including interspecific recombinant congenic strains (IRCS). The originality of this whole model is based on the presence of a small homozygous fragments of *Mus spretus* genome fixed on a *Mus musculus* B6 genetic background [16]. Thus, a strain differs of each other and from the B6 parental strain by the *spretus* segments. *Mus musculus* and *Mus spretus* diverged ~2 million years ago meaning that the association of their two genomes has the potential to lead to genetic incompatibilities [14]. Using this model, in past studies, we were able to localize various QTL modulating male and female reproductive processes. We identified a QTL responsible for ~25% of the embryonic resorptions present in the IRCS-66H-MMU1 strain containing a solely chromosomal fragment of *spretus* origin located on MMU1. This QTL was called *Int2* and has been mapped to an interval of 32 Mb which contains 215 genes [15]. The aim of the present study was to redefine this region and to identify candidate genes potentially involved in embryonic lethality.

To accomplish the fine mapping of this QTL, we generated recombinant substrains from 66H-MMU1 by backcrosses, each of them presenting a unique sub-fragment of the *Int2* QTL. Each recombinant substrain females were crossed with B6 males, resulting in a fetus/placenta complex with heterozygous B6/SEG genes (at the *Int2* locus) and uterine homozygous *spretus* genes (at the *Int2* locus). During each gestation, the substrains were phenotyped *in vivo* by ultrasonography. This non-invasive technology, based upon a high frequency ultrasound device [29] allows *in vivo* real time high resolution observations of embryonic development [15,30] and resorption (~70 µm and ~90 µm lateral and axial resolution, respectively) and permits to carry out longitudinal analysis of gestation. We observed an increase of the embryonic death rate in R3 and R5 substrains (Group 2). The analysis of the genotype/phenotype segregation allowed us to determine two reduced QTL regions (*Int2subA* and *Int2subB*) of approximately 6 Mb each, present together in *spretus* version only in R3 and R5 strains. In the other recombinant substrains which have not the phenotype, the one or the other of the region is present but not the two regions together. So we defined the first reduced *spretus* region called *Int2subA* which encompasses D1Mit50 to D1Mit305 region (>84.5 Mb to 90.5 Mb) and the second called *Int2subB* located at the rs3692309 marker (>92.5 Mb to <100.3 Mb). Our statistical analysis succeeded in proving the presence of *Int2subA* QTL responsible for a main effect on embryonic death but it failed it for *Int2subB*. However, notable differences between the embryonic death rates of certain strains (R6 compared to R3 or R5) led to suppose that this latter region could also have a small effect in the phenotype. Taken together, these data did not support the presence of an epistatic interaction between *Int2subA* and *Int2subB*.

Reverse crosses using IRCS Group 2 males and B6 females revealed that the genes expressed at heterozygous state in the placental tissues are not deleterious for the gestation. Therefore, we deduced that the high rate of embryonic death occurring during the gestation resulted from dysfunction of genes expressed in the uterine tissue. This is in accordance with the normal embryonic development observed in group 2 IRCS females. Then, we carried out a microarray analysis searching to identify uterine deregulated genes in IRCS animals from Group 2. Although we observed deregulated genes located in all chromosomes (19%) we noticed that those situated on the *spretus* fragment were preferentially modified (~40%). This concentration of deregulated genes located on the *spretus* fragment has already been reported in a previous study of our group performed on testis transcriptome [31]. It has been showed that at genomic scale differential SNPs between *Mus musculus* and *Mus spretus* are frequent since they appear, in average, every 100 bp. When located on the promoter regions of *spretus* origin, these nucleotide substitutions could modify the transactivation/transrepression properties of transcription factors of C57BL/6J nature, thus modifying the *spretus* gene expressions. Additionally, dysfunctions leading to embryonic death could result from non-synonymous coding polymorphisms, accumulated during evolution in the *spretus* genome. These phenomena should be originated from evolution of separated genomic regions that produces transcription factors/DNA ("transcriptomic shock") and/or protein-protein ("proteomic shock") incompatibilities [32].

Table 5. Genes of *Led2min* region expressed in uterus (>500 AUF) and displaying a deregulation (R3/B6 ratio) and/or non synonymous SNP.

Gene symbol	Expression level in B6 uterus (AUF)	SNP number	Ratio of expression R3/B6
Led2minA			
Tdp12	12006	4	0.77
Rbm8a	662	6	0.12
AS3MT	521	0	0.28
Sp110	18202	ND	0.52
SP100	8798	19	0.55
Cab39l	1600	0	45
Ankle	13289	1	0.44
Spata3	906	7	1.88
2610045M01Rik	1922	4	0.24
Pcmd1	84109	2	0.90
Hbs2b	1982	4	0.13
Nol	1655	5	0.16
Rdm5d	9972	0	0.15
Capn7b	12526	1	0.62
Eif4a2	662	1	0.50
Ngef	1627	4	0.54
Ltp40	2154	7	0.21
Glpd	517	9	0.11
Arlec	10043	0	0.37
Led2minB			
Colec3	20317	28	2.01
Mlph	8588	10	0.48
Rab17	2599	4	0.34
Lrrfip1	14888	14	1.2
Nomp1	13416	7+1 stop lost	0.78
Rbm	20116	2+1 splice site	0.78
Hmx	12151	1	0.57
Per2	8786	10	0.89
Tnfrsf1	814	9	0.66
Alb1	4667	1	0.57
Ndufs10	52615	2	0.87
Myeov2	14896	5	0.78
Gpc1	1912	1	0.8
Rhpn1	960	1	0.69
Snrnp1	4230	9	3.51
Mterfd2	34052	7	1.32
Ppp1r7	6530	6	0.63
Ndfip	25050	2	1.06
Fcgr2	634	20	0.74
Zfyve	2169	1	0.77
Atg4b	676	8	0.72
Dtymk	9044	1	0.41
Dhrs4	945	3+1 splice site	0.44

ND: Not determined.

doi:10.1371/journal.pone.0043356.t005

Focusing on genes of the *Ld2wisi4* QTL and applying filters from bioinformatics databases, bibliography and our own results, we propose a selection of 7 genes (*Trip12*, *Cab39*, *Psmd1*, *Ncl*, *Cops7b*, *Eif4e2* and *Usp40*) as putative actors of the embryonic death. These genes play a role in VEGF signaling, mTOR signaling and ubiquitine/proteasome-protein degradation pathway. Their effects could be reinforced by a small participation of genes situated on *Ld2wisi8* region and which could act in the same signaling pathways (*Asb1*, *Tuft1p1*, *Ramp1* and *Col6a3*).

Trip12, *Psmd1*, *Cops7b* and *Usp40* from *Ld2wisi4* and *Asb1* from *Ld2wisi8* are involved in protein degradation process through the ubiquitin-proteasome pathway. *Trip12* exerts a ligase activity related to ubiquitination [33]. *Usp40* functions as a deubiquitination enzyme in the same degradation pathway [34] and *Asb1* is a member of the ankyrin repeat and SOCS box (ASB) family. These family proteins interact with Cul5-Rbx2 to form E3 ubiquitin ligase [35]. *Psmd1* is a component of the 26S proteasome. Cops7b is a subunit of the eight-subunit heteromeric Cop9 signalosome complex. Genetic invalidations of some *Cop9* subunits have been associated with developmental defects of post-implantation embryos [36,37,38]. Additionally, *Usp40* functions as a deubiquitination enzyme in the same degradation pathway [34]. *Asb1* is a member of the ankyrin repeat and SOCS box (ASB) family. These family proteins interact with Cul5-Rbx2 to form E3 ubiquitin ligase [35].

In the same manner, *Ld2wisi4 Ncl* gene and *Ld2wisi8 Ramp1* and *Col6a3* genes are involved in angiogenesis. *Ncl* encodes nucleolin and treatment of endothelial cells with anti-nucleolin antibody induces apoptosis of these cells [39]. Moreover, nucleolin associates with VEGF-C62 and can be potentially involved in epithelial cell adhesion and proliferation [39,40]. Concerning *Ramp1* gene, RAMP1 (receptor activity modifying protein) forms a functional receptor for CALCA (Calcitonin gene-related peptide) which is a proangiogenic growth factor in the human placental development and plays a critical role in embryonic development and fetal growth [41]. Concerning the *Collagen type VI alpha 3* gene, COL IV is a main endometrial extracellular matrix component, and an abnormal increased deposition of collagen might impair uterine function, possibly by interfering with vascularization or retarding remodeling events at implantation [42].

Finally, *Cab39* (also called *Mo25*) and *Eif4e2* from *Ld2wisi4* and *Tuft1p1* from *Ld2wisi8*, participate in the mTOR signaling pathway, a regulatory step of protein synthesis and growth. *Cab39* effect has been described upstream of mTOR activation while *Eif4e2* is a downstream signaling target involved in translation initiation. Interestingly, *Mtor* genetic disruption in mice leads to early embryonic death [43]. Homozygous *Tuft1p1* (Tumor necrosis factor alpha receptor 3 interacting protein 1) mutant mice are not viable, *Tuft1p1* mutant mouse line was generated and the enlarged mutant cell size in culture was associated with elevated basal mTOR pathway activity [44]. Otherwise, mTOR pathway is implicated on the VEGF pathway activation. It is worth noting that the VEGF and mTOR pathways have been identified as significantly deregulated in our functional clustering analysis.

Conclusions

We used an *in silico* approach of the embryonic development on a mouse IRCS model to refine a chromosome 1 region (*Ld2*) responsible for embryonic death. The present study succeeded in fine-mapping *Ld2wisi4* QTL which has a main effect on the embryonic death (about 30%) and pointed out a second region *Ld2wisi8* which could have a minor effect on the same phenotype. Collecting and analyzing experimental, bioinformatics and literature data on the expression and function of genes present in the two regions (*Ld2wisi4 and Ld2wisi8*), we propose 7 genes from *Ld2wisi4* that could be related with the phenotype. It appears that the vascularization could be the common denominator at these categories of genes involving angiogenesis and the fluidity of the extracellular matrix. The actual identification of the gene(s) involved in this phenotype will necessarily pass through further molecular approaches. An important outcome of this study is the possibility to evaluate novel promising candidates of RSA in humans [13]. This might contribute to elucidate the molecular basis of this multifactorial and complex human disorder and to propose new diagnostic markers.

Supporting Information

Table S1 Sequences of used real time RT-PCR primers.
(DOC)

Acknowledgments

We thank I. Lanctin and J. Chevalier (Pasteur Institute, Paris) for the IRCS breeding and C. Marchiol (Small animal imaging facility of the Curie Institute, Paris) for their technical assistance. We thank J. Cuquet for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: DV CS AZ XM. Performed the experiments: MV GB GR FM VF. Analyzed the data: MV DV CS AZ PL. Contributed reagents/materials/analysis tools: MV GR FM VF. Wrote the paper: DV CS AZ MV PL.

References

1. Norris W, Nevers T, Sharma S, Kalkunte S (2011) Review: hCG, preeclampsia and regulatory T cells. Placenta 32 Suppl 2: S182–185.
2. Matzuk MM, Lamb DJ (2008) The biology of infertility: research advances and clinical challenges. Nat Med 14: 1197–1213.
3. Zheng K, Yang F, Wang PJ (2009) Regulation of male fertility by X-linked genes. J Androl 31: 74–85.
4. Rai R, Regan L (2006) Recurrent miscarriage. Lancet 368: 601–611.
5. Stephenson MD, Awartani KA, Robinson WP (2002) Cytogenetic analysis of miscarriages from couples with recurrent miscarriage: a case-control study. Hum Reprod 17: 446–451.
6. Philipp T, Philipp K, Reiner A, Beer F, Kalousek DK (2003) Embryoscopic and cytogenetic analysis of 233 missed abortions: factors involved in the pathogenesis of developmental defects of early failed pregnancies. Hum Reprod 18: 1724–1732.
7. Salim R, Regan L, Woelfer B, Backos M, Jurkovic D (2003) A comparative study of the morphology of congenital uterine anomalies in women with and without a history of recurrent first trimester miscarriage. Hum Reprod 18: 162–166.
8. Ray E, Kahn SR, David M, Strauss I (2003) Thrombophilic disorders and fetal loss: a meta-analysis. Lancet 361: 901–908.
9. Levine JS, Branch DW, Rauch J (2002) The antiphospholipid syndrome. N Engl J Med 346: 752–763.
10. Li TC, Makris M, Tomsu M, Tuckerman E, Laird S (2002) Recurrent miscarriage: aetiology, management and prognosis. Hum Reprod Update 8: 463–481.
11. Tulppala M, Palosuo T, Ramsay T, Marttunen A, Salonen R, et al. (1993) A prospective study of 65 couples with a history of recurrent spontaneous abortion: contributing factors and outcome of subsequent pregnancies. Hum Reprod 8: 764–770.
12. Kaare M, Palmer JN, Ulander VM, Kaapa R, Aittomaki K (2006) Variation of the Annexin-like gene is recurrent spontaneous abortions. Mol Hum Reprod 12: 25–29.
13. Mercier E, Lassalle-Lagrave G, Gris JC (2003) JAK2 V617F mutation in unexplained loss of first pregnancy. N Engl J Med 352: 1984–1985.

14. Bidwell BA, Calvert S, Digiovanna A, Georgiou A, D'Hooghe B, et al. (2010) Functional exploration of the adult ovarian granulosa cell tumor-associated somatic FOXL2 mutation p.Cys134Trp (c.402C>G). PLoS One 5: e8380.
15. Lainas P, Bargy G, L'Hote D, Renault G, Marchais-Fournigault C, et al. (2008) Identification of Quantitative Trait Loci responsible for embryonic lethality in mice assessed by ultrasonography. Int J Dev Biol 53: 623–629.
16. Bargy G, Stratzenik M, Gastaut JL, Arnaud MR, Panthier JJ, et al. (2000) Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics 155: 2321–2333.
17. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44–57.
18. Girardi G, Yarkin D, Thurman JM, Holers VM, Salmon JE (2003) Complement activation induces dysregulation of angiogenic factors and causes fetal injection and growth restriction. J Exp Med 203: 2165–2175.
19. Singh J, Ahmed A, Girardi G (2011) Role of complement component C1q in the onset of preeclampsia in mice. Hypertension 58: 716–724.
20. Ahmed A, Singh J, Khan Y, Seshan SV, Girardi G (2010) A new mouse model to explore therapies for preeclampsia. PLoS One 5: e13663.
21. Blois S, Tometsko M, Kandil J, Hagen E, Klapp BF, et al. (2005) Intercellular adhesion molecule-1/LFA-1 cross talk is a proximate mediator capable of disrupting immune integration and tolerance mechanisms at the feto-maternal interface in murine pregnancies. J Immunol 174: 1820–1829.
22. Clark DA, Chaouat G, Arck PC, Mittruecker HW, Levy GA (1998) Cytokine-dependent abortion in CBA × DBA/2 mice is mediated by the procoagulant fgl2 prothrombinase [correction of prothrombinase]. J Immunol 160: 545–549.
23. Bogdarina I, Murphy HC, Burns SP, Clark AJ (2004) Investigation of the role of epigenetic modification of the rat glucokinase gene in fetal programming. Life Sci 74: 1407–1415.
24. Clark DA (2001) Controversies in reproductive immunology. Crit Rev Immunol 21: 215–247.
25. Clark DA, Qualtrough C, Banerjee D, Manuel J, Filipp G (1994) Spontaneous abortion in immunodeficient SCID mice. Am J Reprod Immunol 32: 15–25.
26. Critchley HO, Kelly RW, Lea RG, Drudy TA, Jones RL, et al. (1996) Sex steroid regulation of leukocyte traffic in human decidua. Hum Reprod 11: 2253–2262.
27. Dmowski R (1973) Termination of early pregnancy in rats after ovariectomy is due to immediate collapse of the progesterone-dependent decidua. J Reprod Fertil 35: 185–186.
28. Durlen AJ, Hadshel EK, Barnes MG (2005) Presence of activated macrophages in a murine model of early embryo loss. Am J Reprod Immunol 53: 354–366.
29. Foster FS, Pavlin CJ, Harasiewicz KA, Christopher DA, Turnbull DH (2000) Advances in ultrasound biomicroscopy. Ultrasound Med Biol 26: 1–27.
30. Ouyang YQ, Li SJ, Zhang Q, Cai HB, Chen HP (2009) Interactions between inflammatory and oxidative stress in preeclampsia. Hypertens Pregnancy 28: 56–67.
31. Bradshaw NM, Dudbridge L, L'Hote D, Bernex F, Pigeon J, et al. (2008) Rescue of fertility in homozygous mice for the urokinase plasminogen activator transgene by the transplantation of mouse hepatocytes. Cell Transplant 17: 803–810.
32. Sahin M, Greer PL, Lin MZ, Poucher H, Eberhart J, et al. (2005) Eph-dependent tyrosine phosphorylation of ephexin1 modulates growth cone collapse. Neuron 46: 191–204.
33. Park Y, Yoon SK, Yoon JB (2009) The HECT domain of TRIP12 ubiquitinates substrates of the ubiquitin fusion degradation pathway. J Biol Chem 284: 15440–15449.
34. Quesada V, Diaz-Perales A, Gutierrez-Fernandez A, Garabaya C, Cal S, et al. (2004) Cloning and enzymatic analysis of 22 novel human ubiquitin-specific proteases. Biochem Biophys Res Commun 314: 54–62.
35. Kobirumaki J, Nishiyama T, Nakamura T, Maruhashi Y (2005) ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes. FEBS Lett 579: 6799–6807.
36. Lykke-Andersen K, Schaefer L, Menon S, Deng XW, Miller JB, et al. (2003) Disruption of the COP9 signalosome Csn2 subunit in mice causes deficient cell proliferation, accumulation of p53 and cyclin E, and early embryonic death. Mol Cell Biol 23: 6790–6797.
37. Yan J, Wolz K, Nakamura H, Carattini-Rivera S, Zhao Q, et al. (2003) COP9 signalosome subunit 3 is essential for maintenance of cell proliferation in the mouse embryonic epiblast. Mol Cell Biol 23: 6798–6805.
38. Tomoda K, Yoneda-Kato N, Fukumoto A, Yamanaka S, Kato JY (2004) Multiple functions of Jab1 are required for early embryonic development and growth potential in mice. J Biol Chem 279: 43013–43018.
39. Fogal V, Sugahara KN, Ruoslahti E, Christian S (2009) Cell surface nucleolin antagonism causes endothelial cell apoptosis and normalization of tumor vasculature. Angiogenesis 12: 91–100.
40. Wang ZG, Park TS, Quang RJ (2010) Characterization of novel VEGF (vascular endothelial growth factor)-D splicing isoforms from mouse. Biochem J 428: 343–354.
41. Dong YL, Reddy DM, Green KE, Chauhan MS, Wang HQ, et al. (2007) Calcitonin gene-related peptide (GALCA) is a proangiogenic growth factor in the human placental development. Biol Reprod 76: 890–896.
42. Diao H, Aplin JD, Xiao S, Chen J, Li Z, et al. (2011) Altered spatiotemporal expression of collagen types I, III, IV, and VI in Igsf3-deficient post-implantation mouse uterus. Biol Reprod 84: 255–265.
43. Gangloff YG, Mueller M, Dann SG, Svoboda P, Sticker M, et al. (2004) Disruption of the mouse mTOR gene leads to early postimplantation lethality and prohibits embryonic stem cell development. Mol Cell Biol 24: 9508–9516.
44. Bertani NP, Kin NW, Sharma N, Michaud EJ, Kaminen BA, et al. (2011) Mutations in Trafficking protein particle complex 1 cause specific defects in osteogenesis, craniofacial development, cell growth, and cell size regulation. Dev Biol 360: 66–76.

Article 3

FOXD1 mutations are associated with embryonic resorption in mice and recurrent spontaneous abortion in humans

(Article en préparation)

Résumé

L'objet de cette étude est le clonage positionnel du gène présent dans le QTL *Led1* localisé sur le chromosome 13 de souris et responsable d'une augmentation de la létalité embryonnaire dans la lignée IRCS 66HMMU13. Ce travail devrait permettre de proposer de nouveaux gènes candidats responsables de cas d'avortements spontanés à répétition (RSA).

Cette étude a permis d'associer au phénotype observé des variations de séquence du facteur de transcription Foxd1 qui régule l'expression de gènes impliqués dans l'implantation et la placentation. De plus, des mutations dans ce gène sont en partie responsables de la résorption embryonnaire chez la souris.

L'analyse génétique et fonctionnelle de *FOXD1* chez les patientes RSA a révélé des mutations associées à la pathogenèse. En effet, la mutation p.429AlaAla dans le gène survient dans un pourcentage significatif de cas de RSA. Ces résultats démontrent que *Foxd1* est un nouvel acteur clé de la reproduction des mammifères.

1

FOXD1 mutations are associated with embryonic resorption in mice and recurrent spontaneous abortion in humans

Paul Laissue,[1,2,3] Besma Lakhal,[4*] Magalie Vatin,[1,2*] Julie Cocquet,[1,2] Frank Batista,[5] Gaëtan Burgio,[6,7] Eric Mercier,[8] Esther Dos Santos,[9] Diego Ojeda,[3] Christophe Buffat,[10] Gilles Renault,[1,2] Xavier Montagutelli,[6] Jane Salmon,[11] Philippe Monget,[12] Reiner A. Veitia,[1,2] Marc Fellous,[1,2] Jean-Christophe Gris,[8] and Daniel Vaiman[1,2]

[1]Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Paris, France.

[2]Inserm, U1016, Paris, France.

[3]Unidad de Genética, Facultad de Medicina, Universidad Del Rosario, Bogotá, Colombia.

[4]Department of Cytogenetics and Reproductive Biology, Farhat Hached University Teaching Hospital, Sousse, Tunisia.

[5]Department of Cell Biology, Albert Einstein College of Medicine, New York, NY 10461, USA.

[6]Institut Pasteur, Unité de Génétique des Mammifères, France.

[7]Department of Genetics, Menzies Research Institute, University of Tasmania, Private bag 109, Hobart, 7005, TAS, Australia.

[8]Department of Haematology, University Hospital Nimes, Faculty of pharmacy and Research Team EA 2992, University of Montpellier, France.

[9]Université Versailles-St Quentin, Service de Biochimie et Biologie Moléculaire, UPRES-EA 2493, Faculté de Médecine Paris-Ile de France Ouest, Centre Hospitalier de Poissy, 78303 Poissy Cedex, France

[10]Division of Neonatology, Hôpital la Conception, Assistance Publique-Hôpitaux de Marseille, 147, Boulevard Baille, 13385 Marseille, France.

[11]Department of Medicine, Weill Medical College of Cornell University, New York, NY 10021, USA.

[12] INRA – CNRS - Université de Tours - Haras Nationaux, IFR 135, 37380 Nouzilly, France.

* These two authors contributed equally to the experimental work of this study.

Corresponding author: Daniel Vaiman

Institut Cochin, Université Paris Descartes, Paris, France.

U1016 INSERM-UMR 8104 CNRS

Address:

Faculté de Médecine, Hôpital Cochin

24 rue du Faubourg St-Jacques

75014, Paris, France

Tel :+0033144412301

Fax :+0033144412302

email: daniel.vaiman@inserm.fr

The authors declare that no conflict of interest exists.

Abstract

Human infertility represents a public health concern affecting 10-15% of couples. Despite advances in diagnosis and treatments, ~30% of the cases are considered as idiopathic which suggests a genetic origin of the disease. Recurrent Spontaneous Abortion (RSA) is a frequent cause of infertility. Recently, we localized three Quantitative Trait Loci (QTL) of embryo lethality in a mouse model of interspecific recombinant congenic strains.

Here, we describe the positional cloning of one of these QTL, located on murine chromosome 13. We found that the phenotype is associated with sequence variants of the *Foxd1* transcription factor which regulate the expression of genes involved in implantation and placentation, thus presumably regulating litter size in mice. The genetic and functional analysis of *FOXD1* in RSA patients revealed mutations associated with the disease pathogenesis. In both species, FOXD1 mutations modify its transcriptional capacities on key molecules implicated in pregnancy.

Our results demonstrate that *FOXD1* is a new key actor of mammalian reproduction. Finally, we propose that, in humans, the FOXD1 p.429AlaAla mutation, occurring in a significant percentage of RSA cases, might be used as a molecular marker of the disease.

4

Introduction

Human infertility represents a public health concern affecting 10-15% of couples (1). Recurrent Spontaneous Abortion (RSA), clinically defined by at least three pregnancy losses prior to the 20th week of gestation, is a common cause of infertility, since it affects 1% of pregnancies (2). Among these, 50% are idiopathic which suggest a genetic origin of the disease. Previous attempts to identify RSA causative genes have been relatively unsuccessful due to the intricate interaction of genetic, epigenetic and environmental factors. In 2007, *MK2* was associated with the disease etiology since the p.Val617Phe mutation was encountered in ~1% of the RSA cases increasing the Relative Risk ~5 fold (3). Recently, genome-wide scans have been reported but they did not reach the classical accepted statistical threshold for significance, and apparently failed to identify specific genes (4,5). Genetic association studies of angiogenesis and vasoconstriction-related genes (e.g. VEGF, p53, and eNOS) led to propose some SNPs potentially related to RSA pathogenesis (6).

Here, in order to identify novel RSA etiological genes we opted for a mouse model of Interspecific Recombinant Congenic Strains (IRCS) presenting a clear phenotype of embryonic resorption (7). The IRCS model originates from an initial interspecific cross between *Mus musculus* C57BL/6J and *Mus spretus* SEG/Pas (8). These two mouse subspecies diverged about 1.5 million years ago and exhibit a high genetic and phenotypic polymorphism (9). In the IRCS model, incompatibilities between separate chromosome regions induce strong alterations of gene expression (10). IRCS animals allowed the localization of QTLs responsible for hypofertility and other phenotypes (7, 8, 11, 12). Using this model, we performed in a previous investigation, a QTL (Quantitative Trait Loci) mapping study, aiming at locating genes involved in embryonic resorption and lethality (7).

In the present study, using a combination of positional cloning approaches, transcriptomic analysis, and cellular *in vitro* assays, we obtained data which demonstrate a specific function

of the Foxd1 transcription factor in murine embryonic resorption. Then, by sequencing a collection of idiopathic RSA and control women, we identified *FOXD1* sequence variations which are present exclusively in patients. The functional impact of mouse and human mutations was evaluated in a cell model by assessing modifications of FOXD1 transactivation properties.

Taken together, our results reveal new insights on the molecular aspects of mammalian implantation and early embryonic development. Furthermore, they describe *FOXD1* causative mutations of RSA. Finally, we propose that the most frequent mutation identified in RSA patients (p.429AlaAla) might be used, in combination with other polymorphisms once they are found, as molecular markers of the disease.

Results and Discussion

In a previous work, the *in vivo* exploration of embryonic resorption in IRCS permitted to locate potential causative genes to chromosome segments on mouse MMU1, MMU13 and MMU19 chromosomes (7). The 66H-MMU13 IRCS strain showed a high incidence of embryonic resorption (3-fold increase). This strain contains a unique 2.6Mb *spretus* fragment on MMU13 encompassing 31 genes (7).

In the present study, in order to evaluate placental and uterine transcriptional differences between 66H-MMU13 IRCS and control animals that could explain the phenotype, we performed a cDNA microarray assay in these tissues. We observed a massive deregulation (activation and inhibition, **Supplementary Table 1**) of genes located on several chromosome regions, often located in clusters. Then, we analyzed statistically the putative Transcription Factor Binding Sites (TFBS) present in the promoters of the 50 most upregulated *versus* the 40 most downregulated genes. The putative TFBS were classified according to their degree of discrepancy (in terms of frequency) between both groups of promoters. In placenta, the *Forkhead* (FKHD) binding sites appeared at the top of the list as statistically discrepant between promoters of repressed *versus* induced genes. More specifically, repressed genes contain twice as much FKHD binding sites than the promoters of induced genes (2.54 *vs* 1.27, $p = 0.0009$, **Supplementary Table 1, and Supplementary Figure 1**). Interestingly, *Foxd1* is the sole gene belonging to the *FOX* family located on the MMU13 fragment of *spretus* origin. *Foxd1* was not modified at the transcriptional level. *Foxd1* coding region sequencing of 66H-MMU13 and *spretus* animals revealed 5 non-synonymous variants, p.Asp73Glu, p.Asn126Glu, p.Thr152Ala, p.Asp76_Leu77InsAsp, p.Pro319del, relative to the C57B6/J version. *In silico* analysis underlined the strong interest of one of them, p.Thr152Ala, since it affects the *forkhead* DNA-binding domain at a highly conserved position among mammalian

species (**Figure 1**) and the SIFT software predicts a deleterious effect. Additionally, this mutation triggers the disappearance of a putative hot-spot of phosphorylation.

Then, in order to investigate the transactivation capacity on target genes of distinct Foxd1 versions (*musculus*, *spretus* and p.Thr152Ala) we performed a luciferase reporter assay using promoter constructs of two genes (*Pgf* and *C3*) which were strongly modified in the microarray experiment and display key roles in mammalian implantation. Indeed, the Placenta Growth Factor transcript (*Pgf*, a validated Foxd1 target gene), was 4.3 fold less abundant in the 66H-MMU13 uterus compared to the C57BL/6J control; meanwhile complement component C3 was up-regulated 5.7 fold in the 66H placenta compared to C57BL/6J control (13). For this purpose, we cloned *Foxd1*-C57BL/6J (Foxd1-mm) and *Foxd1*-SEG/Pas (Foxd1-seg) coding regions into expression vectors, in order to test the capacity of Foxd1 to transactivate the *Mus Musculus* and *Mus spretus Pgf* promoters (Pgf-mm and Pgf-seg, respectively, **Figure 2a**). We found that even though Foxd1-seg was able to stimulate the expression of Pgf-mm, the activation was strongly and significantly altered (~3 fold reduction) compared to that of Foxd1-mm on the Pgf-mm promoter, as in the transcriptome assays. Reciprocally, we observed that the Foxd1-mm transactivation capacity on the Pgf-seg promoter was reduced compared to its effects on the Pgf-mm promoter. This could be due to the natural genomic polymorphism, in terms of nucleotide substitutions, observed between *Musculus* and *Spretus* genomes (10). Interestingly, our experiments showed that the Foxd1-mm version is significantly more effective to transactivate the Pgf-seg promoter than the Foxd1-ms version itself. Hypothetically, this is consistent with the idea that the *Mus spretus Foxd1* variants modulate the number of progeny to an optimal (and not maximal) number. On this point, it is worth noting that the *spretus* litter size is naturally smaller than the *musculus* one (5.3 ± 1.4, *versus* ~7.6 in C57BL/6J) (14, 15). We also tested the effect of the Foxd1-T152A mutation alone on the expression of Pgf-mm. As expected, this mutation significantly

(albeit moderately) decreased the transactivation capacity of Foxd1 on the *Pgf* promoter contributing to the downregulation of the *Pgf* levels observed in the 66H-MMU13 uterine tissue, indicating nevertheless that the other Foxd1 *spretus* variants also participate in this phenomenon (**Figure 2b**). Then, we investigated the functional impact of *spretus* Foxd1 on the *Mus musculus* complement *C3*, a gene playing a central role in embryonic resorption, as shown by the CBA x DBA cross. Foxd1-Sep overexpression induces higher expression levels of C3-mm than those induced by the Foxd1-mm version (**Figure 2c**) while the isolated Foxd1-T152A mutation did not show significant differences from the control on this promoter suggesting that the other Foxd1 variants should contribute to the phenotype. These *in vitro* experiments are consistent with the transcriptomic data in which C3 was 5.7 fold more abundant in the 66H-MMU13 placenta than in the C57BL/6J tissue. Our experiments (transient transfections) also suggest that C3 is a direct target of Foxd1, an observation that was not reported before.

In sum, Foxd1 regulates the expression of two crucial (*Pgf* and *C3*) genes implicated in pregnancy maintenance and its mutations in mouse are related to embryonic resoption. *Pgf* is highly expressed in the placenta, where it regulates the differentiation of vascular endothelia (16). The uterine NK cells (uNK), an endometrial lymphocyte cell subset transiently found during endometrial decidualization and essential for immune dialogue with trophoblasts, express *Pgf* and *Vegf* (17,18,19). $Pgf^{-/-}$ mice display defects in early differentiation of binucleate uNK cells (20). The links of *C3* with pregnancy are abundantly described in literature: its activation is required for an antiphospholipid-induced pregnancy loss that can be reverted by the administration of heparin which blocks the activation of the complement cascade (21). In normal pregnancy, the placenta appears to be subjected, at the materno-foetal interface, to a complement-mediated immune attack. Thus, an appropriate complement inhibition is required for a physiological gestation and, as it has been thoroughly

demonstrated in mice, the deficiency of complement regulatory proteins progressively leads to the embryonic lethality (22, 23). Indeed, excessive local complement C3 production may overwhelm the complement regulatory mechanisms leading to pregnancy loss. These findings suggest that strong disturbances of C3 expression (activation or inhibition) are related to embryo abortion. We attempted to formerly demonstrate the role of *Foxd1* by creating a *Mus musculus* knock-in mouse in which the *musculus* version would be substituted by the *spretus* version. For this purpose, we generated ES cells in a 129 genetic background. However, chimaeric mice obtained from these cells always exhibited a low chimerism level, and were very difficult to obtain in significant amounts. This suggests that Foxd1 *spretus* might have a negative effect on implantation, which leads to a counter-selection of relevant founder mice.

Then, in order to assess whether mutations in this gene would generate a similar phenotype in the human species we sequenced its entire open reading frame in a panel of 556 idiopathic RSA and in 271 women with normal fertility. We found a total of 27 distinct variants (**Table 1**). Eighteen were found only in RSA patients (all at heterozygous state). Mutations found in RSA women concern 33 patients (5.9%). The statistical comparison by a contingency Chi 2 test between the exclusive variants identified in the patient population *versus* those present in the control group was highly significant ($p = 0.00056$). The relative risk, calculated using all these 'private' mutations, was 10.3 (5% confidence interval: [1.4 – 77.2]). Interestingly, the statistical significance was achieved using a unique sequence variant (p.429AlaAla) in which two alanine residues were inserted at position 429 of the protein. This mutation was found 12 times exclusively in the patient group (2.2%, $p = 0.015$). It seems also that the p.Ala88Gly may have a protective effect since it is frequently encountered in the control group (**Table 1**).

10

Next, we assessed the functional effects of three human *FOXD1* variants predicted to be damaging as they affect nucleotide positions conserved among mammalian species. These variants were exclusively found in RSA women and were not previously described in SNP databases. For this purpose, we performed similar functional tests that those carried out for *Foxd1* variants. We tested the FOXD1 wild type (WT) and mutant versions (FOXD1-Ala356Gly, FOXD1-Ile364Met and FOXD1-429AlaAla) for their transactivation properties on *PGF* and *C3* promoters (**Figure 3a and 3b**). The FOXD1-Ala356Gly and FOXD1-429AlaAla were completely unable to activate the PGF promoter, while the Ile364Met retained a transactivation capacity. These results suggest that variants of FOXD1 may reduce PGF levels, thus potentially affecting pregnancy maintenance.

Concerning the C3 promoter, the FOXD1-WT version induced C3 expression while the FOXD1-Ala356Gly and FOXD1-Ile364Met versions had no effect. Conversely, the FOXD1-429AlaAla construct strongly induced the transactivation of the C3 promoter. Therefore, the *FOXD1* p.429AlaAla should be considered as a hypomorphic and hypermorphic allele on PGF and C3 promoters, respectively. In humans, it has been shown that C3 levels are higher in patients presenting a third consecutive pregnancy loss than in women with a successful pregnancy after two idiopathic abortions (24). In mice, moderate levels of C3 expression are necessary for a successful gestation, since at 15 days *post coitum*, C3 knock-out (KO) animals display a much higher resorption rate than their WT counterparts, due to dysfunctions of trophoblast and labyrinth development (25). Our *in vitro* experiments suggest that *FOXD1* mutations that decrease (to a basal level) or increase C3 expression are deleterious. This indicates that a subtle tuning of C3 expression levels is necessary for an optimal function of the foeto-placental interface.

We have to note that, up to now, *Foxd1* KO animals have essentially shown (only at homozygous state) severe developmental defects of kidney and optic chiasm development

(26, 27). The *FOXD1* human mutations (or the mouse Foxd1-ms substitution) described here display milder functional effects than those observed in *Foxd1* homozygous KO mice. Thus, we hypothesize that more drastic mutations in human may be related to kidney and/or neurological phenotypes. In this context, we cannot discard that *FOXD1* should be implicated in syndromic forms of RSA which include these and/or other clinical features.

In conclusion, we have validated for the first time the positional cloning of a QTL by genotyping an outbred species (human). The approach outlined may be useful for other positional cloning projects, where a limited set of recombination events are available for mapping purposes (as in the IRCS, or more generally in mouse strains). In addition, our results show that *FOXD1* is a new key molecular actor modulating pregnancy maintenance. Its mutations are partly responsible for embryonic resorption in mice and RSA in humans. Finally, we propose that the FOXD1 p.429AlaAla mutation can be considered as a RSA molecular marker which can be easily tested by PCR/sequencing.

Methods

Human subjects

RSA patient group consisted of 556 women enrolled in a matched case-control study of unexplained pregnancy loss in the Nimes Obstetricians and Haematologists study of at least three consecutive abortions (n= 2175 patients) (28). The control group consisted of 271 women each of whom had at least one live birth with no history of pregnancy loss. All these subjects (RSA patients and controls) are from Caucasian origin.

Information on expression microarrays, bioinformatics, sequence analysis plasmid constructs and cell culture assays, can be found in the Supplemental Methods.

Study approval

The institutional review board of the Hospital of Nimes approved all steps of clinical the study. All individuals gave written informed consent. Experiments with animals were in accordance with the policies of the Paris Descartes University, the Cochin Institute and the Guidelines for Biomedical Research Involving Animals.

Acknowledgements: This work has been partly funded by the Grant "MAMMIFERT" from the Agence Nationale de la Recherche (ANR). PL has been supported by grants from the Roche Research Foundation and the ANR. Expert sequencing by the sequencing platform of the Cochin Institute is greatly acknowledged.

References:

1. Feng HL. Molecular biology of male infertility. *Arch Androl.* 2003;49(1):19-27.

2. Rai R, Regan L. Recurrent miscarriage. *Lancet.* 2006;368(9535):601-611.

3. Mercier E, Lissalde-Lavigne G, Gris JC. JAK2 V617F mutation in unexplained loss of first pregnancy. *N Engl J Med.* 2007;357(19):1984-1985.

4. Kotte AM, et al. A genome-wide scan in affected sibling pairs with idiopathic recurrent miscarriage suggests genetic linkage. *Mol Hum Reprod. 2011*;17(6):379-385.

5. Li W, Zeng Chan W, Cui X, Xiao Feng L, Mao Sheng Y. Genome-wide screening for risk Loci of idiopathic recurrent miscarriage in a Han Chinese population: a pilot study. *Reprod Sci.* 2010;17(6):578-584.

6. Su MT, Lin SH, Chen YC. Genetic association studies of angiogenesis- and vasoconstriction-related genes in women with recurrent pregnancy loss: a systematic review and meta-analysis. *Hum Reprod Update.*2011;17(6):803-812.

7. Laissue P, et al. Identification of Quantitative Trait Loci responsible for embryonic lethality in mice assessed by ultrasonography. *Int J Dev Biol.*2009;53(4):623-629.

8. Burgio G, Szatanik M, Guenet JL, Arnau MR, Panthier JJ, Montagutelli X. Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. *Genetics.* 2007;177(4):2321-2333.

9. Dejager L, Libert C, Montagutelli X. Thirty years of Mus spretus: a promising future. *Trends Genet.* 2009;25(5):234-241.

10. L'Hôte D, Serres C, Veitia RA, Montagutelli X, Ouimouden A, Vaiman D. Gene expression regulation in the context of mouse interspecific mosaic genomes. *Genome Biol.* 2008;9(8):R133.

11. L'Hôte D, et al. Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. *Genetics.*2007;176(3):1907-1921.

12. L'Hôte D, et al. Fidgetin-like1 is a strong candidate for a dynamic impairment of male meiosis leading to reduced testis weight in mice. *PLoS One.* 2011;6(11):e27582.

13. Zhang H, et al. Transcriptional activation of placental growth factor by the forkhead/winged helix transcription factor FoxD1. *Curr Biol.*2003;13(18):1625-1629.

14. Vargas MJ, Palomo LJ, Palmqvist P. Reproduction of the Alegrian mouse (*Mus spretus* Lataste, 1883) in the south of the Iberian Peninsula. *Bonn. zool. Beitr.*1991; 42:1-10.

15. Casellas J, Medrano JZ. Within-generation mutation variance for litter size in inbred mice. *Genetics.*2008;179 (4):2147-2155.

16. Persico MG, Vincenti V, DiPalma T. Structure, expression and receptor-binding properties of placenta growth factor (PlGF). *Curr Top Microbiol Immunol.* 1999;237:31-40.

17. Clark DE, Smith SK, Licence D, Evans AL, Charnock-Jones DS. Comparison of expression patterns for placenta growth factor, vascular endothelial growth factor (VEGF), VEGF-B and VEGF-C in the human placenta throughout gestation. *J Endocrinol.*1998;159(3):459-467.

18. Li XF, et al. Angiogenic growth factor messenger ribonucleic acids in uterine natural killer cells. *J Clin Endocrinol Metab.*2001;86(4):1823-1834.

19. Lash, G.E., Schiessl, B., Kirkley, M., Innes, B.A., Cooper, A., Searle, R.F., Robson, S.C., and Bulmer, J.N. 2006. Expression of angiogenic growth factors by uterine natural killer cells during early pregnancy. *J Leukoc Biol* 80(3):572-580.

20. Tayade C, et al. Genetic deletion of placenta growth factor in mice alters uterine NK cells. *J Immunol*.2007;178(7):4267-4275.

21. Girardi G, Redecha P, Salmon JE. Heparin prevents antiphospholipid antibody-induced fetal loss by inhibiting complement activation. *Nat Med*.2004;10(11):1222-1226.

22. Xu C, Mao D, Holers VM, Palanca B, Cheng AM, Molina H. A critical role for murine complement regulator crry in fetomaternal tolerance. *Science*.2000;287(5452):498-501.

23. Molina H. Complement regulation during pregnancy. *Immunol Res*.2005;32(1-3):187-192.

24. Sugiura-Ogasawara M, Nozawa K, Nakanishi T, Hattori Y, Ozaki Y. Complement as a predictor of further miscarriage in couples with recurrent miscarriages. *Hum Reprod*.2006;21(10):2711-2714.

25. Chow WN, Lee YL, Wong PC, Chung MK, Lee KF, Yeung WS. Complement 3 deficiency impairs early pregnancy in mice. *Mol Reprod Dev*.2009;76(7):647-655.

26. Hatini V, Huh SO, Herzlinger D, Soares VC, Lai E. Essential role of stromal mesenchyme in kidney morphogenesis revealed by targeted disruption of Winged Helix transcription factor BF-2. *Genes Dev*.1996;10(12):1467-1478.

27. Herrera E, et al. Foxd1 is required for proper formation of the optic chiasm. *Development*.2004;131(22):5727-5739.

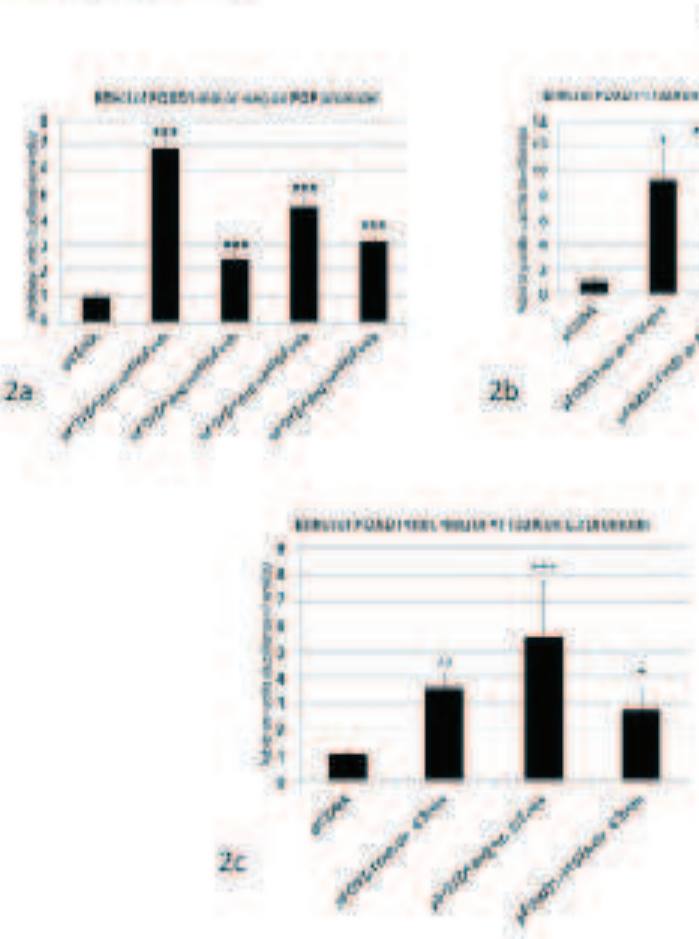
28. Lissalde-Lavigne G, et al. Factor V Leiden and prothrombin G20210A polymorphisms as risk factors for miscarriage during a first intended pregnancy: the matched case-control 'NOHA first study. *J Thromb Haemost* 2005; 3(10):2178-2184.

17

Figure 1: Alignments of FOXD1 DNA binding domain in different species of vertebrates. Note the difference between the *sprattus*-derived samples and all the other species at position 152.

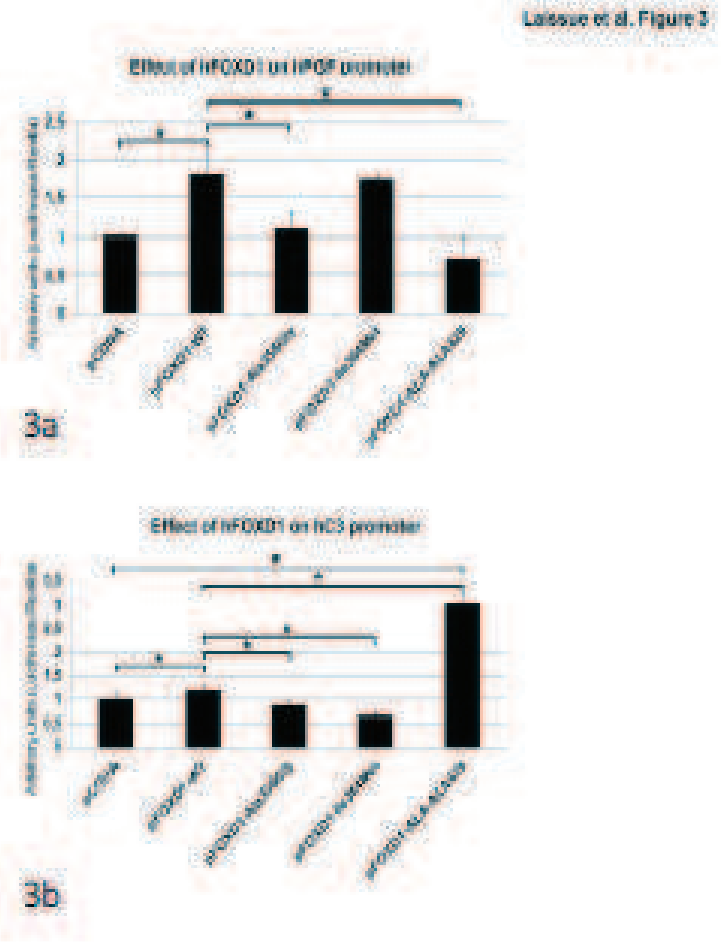
18

Figure 2: Transactivation properties of distinct mouse Foxd1 versions on *Pgf* and *C3* promoters. **2a, 2b)** Effects on *Pgf* promoter. In 2a) and 2c) the stars correspond to post-hoc tests of each category relative to pcDNA. **2c)** Effects on *C3* promoter. **mm**: *mus musculus*. **seg**: *mus spretus*. **pCDNA**: Empty vector. On the graphs errors are represented as standard errors to the mean (S.E.M).



19

Figure 3: Transactivation properties of human WT and mutant FOXD1 versions on *PGF* and *C3* promoters. **3a)** Effects on *PGF* promoter. **3b)** Effects on *C3* promoter. **pCDNA**: Empty vector. On the graphs errors are represented as standard errors to the mean (S.E.M).



20

Table 1. *FOXD1* open reading frame sequencing in RSA patients and control individuals. Mutations tested for their functional impact are marked in bold characters.

DNA	Protein	Patients (n=556)	Controls (n=271)	P values < 0.05
c.69G>C	p.Gly23Gly	7	3	
c.237G>A	p.Leu79Leu	3	0	
c.300C>T	p.Ala100Ala	1	0	
c.324 C>T	p.Gly108Gly	2	0	
c.612G>A	p.Glu204Glu	1	2	
c.903C>A	p.Ala301Ala	1	0	
c.1007 C>T	p.Ala336Val	1	0	
c.1248G>C	p.Val416Val	1	0	
c.1297 GCC>GCG	p.Ala432Ala	2	0	
c.1308A>G	p.Ser436Ser	318	136	
c.1055 C>G	p.Arg352Pro	1	0	
c.1007 C>T	p.Ala336Val	1	0	
c.164G>C	p.Arg55Pro	1	0	
c.263G>C	p.Ala88Gly	20	23	0,003796034
c.326_327InsGCG	p.Ins109Gly	1	0	
c.683C>T	p.Pro228Leu	0	2	0,042799072
c.721G>C	p.Ala241Pro	1	0	
c.976G>A	p.Ala326Thr	21	8	
c.1067 C>G	**p.Ala356Gly**	**1**	**0**	
c.1092 C>G	**p.Ile364Met**	**2**	**0**	
c.909_1165del256	p.FS>STOP376	1	0	
c.1146-1160del	p.Gln383_Ala387del	27	13	
c.1169_1170InsGCCCGC	p.Ins391ProPro	6	7	
c.1187C>T	p.Pro396Leu	3	1	
c.1285_1286InsGCCGCG	**p.Ins429AlaAla**	**12**	**0**	**0,015586607**
c.1309G>A	p.Val437Ile	1	0	
c.1324G>T	p.Ala442Ser	1	0	

Article 4

A frequent polymorphism in ALPP protects from recurrent spontaneous abortions

(Article en préparation)

Résumé

L'analyse de la résorption embryonnaire chez la souris nous a permis d'identifier des polymorphismes du gène *Alpp* de la phosphatase alcaline placentaire comme potentiellement responsables d'une augmentation de la mort embryonnaire chez la souris.

Chez la femme, ce gène exprimé dans le placenta est associé à des pathologies de la grossesse telles que la prématurité, la pré-éclampsie et le retard de croissance intra-utérin. Cependant, peu d'informations sont disponibles concernant le rôle d'ALPP dans les RSA.

Nous avons séquencé ce gène à partir une série de 100 femmes contrôles et 100 patientes RSA. Une analyse fonctionnelle a été effectuée en culture cellulaire. Nous avons montré que plusieurs allèles et combinaisons alléliques sont surreprésentés dans les cas de RSA. En revanche, l'un des polymorphismes non synonymes (ILE89LEU) a été associé à une protection vis-à-vis de la maladie. L'analyse fonctionnelle de ce variant a révélé une augmentation significative de l'activité de la phosphatase. Le génotypage de deux polymorphismes clefs a été entrepris sur un nombre accru de patientes (550 patientes et 250 contrôles).

A frequent functional polymorphism in the placental alkaline phosphatase gene protects from Recurrent Spontaneous Abortions.

Magalie Vatin[#], Sylvie Bouvier[#], Linda Bellazi, Xavier Montagutelli, Paul Laissue, Ahmed Ziyyat, Catherine Serres, Philippe De Mazancourt, Marie-Noelle Dieudonné, Etienne Mornet, Daniel Vaiman, Jean-Christophe Gris

Departement of Haematology, University Hospital, and EA2992, University of Montpellier 1, Nîmes; Laboratory of Haematology, Faculty of Pharmaceutical and Biological Sciences, University of Montpellier 1, Montpellier; FRANCE.

Unité de Pathologie cellulaire et génétique, UPRES-EA2493, UFR Paris Ile de France Ouest PRES Universud Paris, Université de Versailles Saint-Quentin en Yvelines, Versailles, France (LB, PDM, MND,EM)

#Contributed equally.

BACKGROUND: Fertility is a quantitative complex character governed by a considerable number of genes. Despite clinical and scientific advances, several cases of human infertility remain unexplained. The use of an original mouse model created at the Pasteur institute, permit us to identify a candidate gene, *ALPP, encoding* the Placental Alkaline Phosphatase (PLAP), potentially responsible for recurrent miscarriage in women. This gene expressed in the placenta has been reported to be associated with pregnancy diseases such as pre-term delivery, preeclampsia and Intra Uterine Growth Restriction. However, little information is available regarding the role of *ALPP* in recurrent spontaneous abortion (RSA). This study was conducted to investigate the association between polymorphisms of *ALPP* and idiopathic miscarriages.
METHODS: This gene was sequenced in a series of 100 controls and 100 patients affected by Recurrent Spontaneous Abortions, from the same ethnic background.
RESULTS: We showed that several alleles and allelic combination are more frequent in RSA women. In particular, one of the polymorphism induces an amino-acid change in the protein. To better understand the impact of the alteration of the PLAP protein on this function, both the normal and modified versions of the protein were cloned in an expression vector and their activity analyzed. This analysis revealed a significant different in activity between the two variants.

CONCLUSION: This study confirms that mouse models are helpful to decipher complex multifactorial diseases in humans and is a contribution in the understanding of the relation between *ALPP* and RSA.

Key words: PLAP / recurrent pregnancy loss / single nucleotide polymorphism / haplotypes

INTRODUCTION

Fertility is the result of a series of complex interactions between parents, between gametes and in viviparous species, between the placenta and uterus that are hemi-allogenic. This complexity is sustained by a large number of genes, involved in gametogenesis, gamete interaction, fetal and maternal immunity, regulation of tolerance, vasculogenesis and so on (for a review, see (Matzuk and Lamb, 2008)). Imbalance in any factor leads to infertility and more often to reduced fertility. Infertility is a major public health concern, since about 80 million of couples are affected around the world. Miscarriage is presumably the most common complication of pregnancy, and is defined as the spontaneous termination of pregnancy before the fetus has attained the stage of viability. Miscarriage includes all pregnancy losses from the time of conception until 24 weeks of gestation. In human clinic, several causes have been associated with this pathology: uterine anatomic abnormalities, thrombophilic disorders, infective and systemic diseases, immune, endocrine, genetic and epigenetic causes. There are two types of miscarriage: sporadic and recurrent. Spontaneous recurrent miscarriages (SRM), or Recurrent Spontaneous Abortion (RSA) are defined by the occurrence of three or more consecutive pregnancy losses and affect 1% of couples. However, many clinicians define recurrent miscarriage as two or more losses thus increasing the number of affected patients from 1% to 5% (Hogge et al., 2003; Rai and Regan, 2006). After standard gynecological, hormonal and karyotypic investigations the etiology remains unknown in 50% of cases (Jauniaux et al., 2006; Kupferminc, 2003).

To identify genetic causes of this complex pathology, we have used a mouse model of interspecific recombinant congenic strains (IRCS) established at the Pasteur Institute in Paris (Burgio et al., 2007; L'Hote et al., 2007). These strains are the result of an initial cross between two species *Mus spretus* and *Mus musculus* C57BL6/J, two species that diverged

around 1.5 million years ago (L'Hote et al., 2010), followed by two backcrosses and by sister-brother matings for more than 30 generations. All the 53 existing IRCS encompass ~2% of *spretus* genome into a *musculus* background. This model has proven to be a powerful tool for dissecting fertility QTL (L'Hôte et al., 2011). In 2009, phenotyping, using *in vivo* ultrasound biomicroscopy allowed us to identify the 66HMMU1 strain as presenting an increase of embryonic death by embryonic resorption (25% compared to less than 10% in the C57BL6/J control). This phenotype has been associated with the presence of a QTL region of ~32Mb encompassing 215 genes localized on the unique *spretus* fragment of mouse chromosome 1 (Laissue et al., 2009). This results sustained the idea that one or more genes in this fragment, *spretus* in sequence present a certain incompatibility with the *musculus* context, and trigger defects in implantation/early development (L'Hote et al., 2010).

Applying a candidate-gene approach, we focused on one of these genes, the Alkaline phosphatase placental-like 2 (*Alppl2*), due to his placental expression in human and because the knock-out in mouse induce a slower development and higher rates of degeneration during *in vitro* preimplantation development (Dehghani et al., 2000). Furthermore, in humans, previous studies on *ALPP* (encoding the PLAP protein) have revealed an association between some genetic polymorphisms and pathologies occurring during pregnancy in different populations (Northern Europe and Asia). Two polymorphisms P209R and E429G have been associated with spontaneous abortions (Wennberg et al., 1995; Wennberg et al., 2002). Another study revealed a sequence variation in the *ALPP* promoter which could account for associations with pregnancy diseases (Bellazi et al., 2010). Recently, *ALPP* has also been associated with pregnancy diseases, such as pre-term delivery (Ferianec and Linhartova, 2011). In addition it is known that PLAP activity increases throughout pregnancy and

suddenly decreases after delivery (Whyte et al., 1995). However, its physiological role is still poorly understood.

Data from mouse studies and from the bibliography prompted us to study the impact of this gene in spontaneous abortions in a Mediterranean population. In the present work we performed a retrospective case/control study of recurrent miscarriages by analyzing by sequencing a sample of human DNAs from patients and controls. We identified variants of PLAP that are associated with RSA, one of them (Ile89Leu) affecting the coding sequence. Both the normal and modified versions of the protein were cloned in an expression vector and their enzymatic activities were analyzed. A significant difference between the two variants was revealed. This study may contribute to clarify the role of *ALPP* in the RSA phenotype of the patients.

MATERIAL AND METHODS

Patients

A 1:1 retrospective case control study was performed at the University Hospital of Nîmes (France); patients being a series of 100 women with unexplained primary recurrent miscarriages (childless women with two to five pregnancy losses) selected from the Nîmes Obstetricians and Haematologists (NOHA) cohorts. A total of 33 women had only experienced early, embryonic losses (≤ 10 weeks of gestation); 54 had only experienced late, fetal losses (> 10 weeks of gestation) and 13 had experienced both events. These women were compared with 100 control women with at least two children and no miscarriage reported. Characteristics of the study population were summarized in table 1.

The study was approved by the University Hospital of Nîmes Institutional Review Board and ethics committee. This clinical investigation was performed according to the Helsinki declaration of 1975 as revised in 1996. All the women had given their informed consent to participate.

Methods

ALPP was amplified by polymerase chain reaction (PCR) with the Kapa Hifi polymerase (KAPA BIOSYSTEMS) followed by sequencing performed on the human genetic technical platform at the University Hospital of Nîmes, France.

Amplification of *ALPP*

A 4,448 bases pair fragment of *Alpp* (11 introns and 12 exons) was amplified reliably and reproducibly using the forward primer: 5'TCACCCAAATGGCATCTGTACTCAA-3' and the reverse primer 5'-GTTATCCACCCCTCCAGACACTGTG -3' with the annealing temperature

fixed at 66°C. Conditions were the following: DNA 100 ng, primers 10 µM, DMSO 5%, Kapa Hifi ready mix 1X, H20 to 25 µL. PCR program used was the following: pre-denaturation: 98°C (2 minutes), denaturation: 98°C (20 seconds), hybridization: 66°C (15 seconds), elongation: 72°C (150 seconds), 35 cycles, final elongation: 72°C (150 seconds) and 4°C.

Sequencing of *ALPP*

Commercially available reagents and kits were used. Purification of PCR products was performed with the kit Promega (Wizzard®) and then sequenced with the amplification kit BigDye XTerminator® (Applied Biosystems). Reactions were carried out using standard conditions supplied by the company with nine specific primers covering the entire coding sequence, at the annealing temperature of 50°C.

-*Alpp* 578: 5'TCACCCAAATGGCATCTGTA-3', *Alpp*767: 5'- GGGACACAGTTCTCCCTGAT-3', *Alpp* 1322: 5'-AAACTGGGGCCTGAGATACC -3', *Alpp* 1500: 5'-TGGAATCCCAGAGGACAGAG-3', *Alpp* 1985: 5'-CTCCAACATGGACATTGACG -3', *Alpp* 2431: 5'-CAGAGGTGTGGGGCTCAG-3', *Alpp* 2867: 5'-AGGGCAGGCTCAGCATCT -3', *Alpp* 3230: 5'-CTGACCAGGCAAAACGTG-3', *Alpp* 3363: 5'- TCCTATACGGAAACGGTCCA-3'

The sequence analysis was performed through the software Seqscape® (Applied Biosystem).

Alkaline phosphatase activity

A full-length Pl^{l} cDNA of *ALPP* (NM_001632)was cloned in the pcDNA3.1 plasmid (Invitrogen) by standard molecular biology methods. The Ile89Leu variant was obtained directly from the Pl^{l} cloned cDNA by site-directed mutagenesis. The mutated cDNA was

fully sequenced to make sure that the target mutation (rs13026692) was inserted without other mutations. *Pl*^{I} or Ile89Leu plasmids were transiently transfected in COS-7 cells for 48 h as previously described (Fauvert et al., 2009). Empty pcDNA3.1 plasmid was used as negative control. The plasmid pcDNA3.1/His/LacZ containing the β-galactosidase gene was used as a positive control of transfection. Total alkaline phosphatase activity was measured with a COBAS Integra 800 automate (Roche) and was weighted with β-galactosidase activity. For each variant, experiments were repeated six times independently. All values were expressed as medians of the separate experiments and statistical analysis was performed using paired Wilcoxon test.

3D-modelingA 3D model of the PLAP molecule has been previously obtained by using its 1,8 Å-resolution crystal structure (Le Du et al., 2001). The variant discussed here was visualized and analyzed using the open source PyMOL software http://www.pymol.org/.

Statistical methods

Statistical analysis and association with single markers and haplotypes between patients and controls were performed using Statview® (5.0, Abacus Concepts, Berkeley, United States) and Haploview® (4.1 Daley Lab at the Broad Institute, Cambridge, USA) for the study of haplotypes. The Chi-squared test and logistic regression analysis were used to compare the two populations. For all analysis, $p \leq 0.05$ was considered statistically significant. Relative risk was estimated from logistic odds ratio (OR) with a 95% confidence interval (CI).

RESULTS

Genotyping results

49 polymorphisms were identified, 36 in introns and 13 in the coding sequence. Among these polymorphisms, 7 are not synonymous, affecting one or both alleles of the gene.

Univariable data analysis

Univariable and multivariable analysis on these polymorphisms were performed. Single nucleotide polymorphisms (SNPs) and results of univariable analysis are summarized in table 2. Six variants appeared to be significantly different between patients and controls: *Alpp* 988, *Alpp* 989, *Alpp* 1338 (Ile89Leu, rs130226692), *Alpp* 2498 (Tyr268Tyr, rs2260309), *Alpp* 2621 and *Alpp* 3315.

Four variants appeared as risk factors for miscarriages: *Alpp* 988, *Alpp* 989, *Alpp* 2498 (Tyr268Tyr, rs2260309) and *Alpp* 2621 with relative risks of 2.698, 2.698, 2.253, and 3.093, respectively (see confidence intervals in Table 2).

Two variants appeared as protecting from miscarriages: *Alpp* 1338 (Ile89Leu, rs130226692) and *Alpp* 3315, with relative risks of 0.474 and 0.538, respectively (see Table 2 for 95% confidence intervals).

Four SNPs were located in non-coding regions: *Alpp* 988, *Alpp* 989, *Alpp* 2621 and *Alpp* 3315, while two were in the Open Reading Frame; one is a synonymous SNP (Tyr268Tyr, rs2260309) and one is a missense SNP (Ile89Leu, rs130226692).

We attempted to classify the significant SNPs according to the subtypes of recurrent miscarriages (Table 3): no striking predominance of any SNP in a peculiar clinical subgroup could be evidenced.

Multivariable analysis

After stepwise multivariable logistic regression analysis, two polymorphisms were identified to be independently associated with the risk of miscarriages (Table 4): Tyr268Tyr (rs2260309), a favoring SNP (RR = 2.82) and Ile89Leu (rs130226692), a protecting SNP (RR = 0.27). The analysis of these two SNP-derived genotypes found heterozygosity for the Tyr268Tyr (rs2260309) SNP and homozygosity for the Ile89Leu (rs130226692) SNP to be independently associated with the risk of miscarriages, the first one being a favoring genotype, the second one being a protecting genotype (Table 5).

Linkage disequilibrium analysis and haplotype construction

The results of the haplotype study are summarized in the Haploview LD Plot® (Figure 1). The linkage equilibrium coefficient (LD) was calculated for each pair of SNPs and haplotypes were divided into blocks. The comparisons of haplotype frequencies between cases and controls made it possible 10 of them being associated with recurrent pregnancy losses, 5 being positive risk factors, 5 being negative risk factors (Table 6).

We then analyzed the different haplotypes for different combinations (Figure 2), identifying a subset of frequent combinations amongst a theoretical total of 1640 combinations being possible.

We restricted the subsequent statistical analysis to the more frequently evidenced haplotypes (frequencies ranging from 10.83% to 1%) (Table 7): the frequencies of two among these 18 combinations ("B" and "Q") were significantly different in patients compared to controls, two other combinations being very close to the statistical threshold ($p = 0.065$).

From the previous results, univariate analysis (Table 8) then multivariate (Table 9) analysis were performed, entering in the models the four haplotype combinations with a p value lower than 0.07. The two combinations "B" and "Q" were finally identified as independent positive risk factors for miscarriages, the later one being associated with a 5 fold increase of the clinical risk.

Going back to the subtypes of recurrent miscarriages, the "B" combination was found in 2 women belonging to the recurrent embryonic loss subgroup (2 out of 33, 6%) but was absent from the 2 other subgroups. The "Q" combination was positive in 4 women from the embryonic loss subgroup (4 out of 33, 12%) and in 6 women from the fetal loss subgroup, and thus may explain up to 10% of the cases of recurrent miscarriages.

Functional analysis of PLAP variants

We then wished to evaluate the biological effects of the ‘protective’ Ile89Leu variant. For this we cloned the two versions of the gene in an expression vector. Then, after transfection in COS-7 cells,the alkaline phosphatase activity was measured. COS-7 cells do not express endogenous alkaline phosphatase. The enzymatic activity was increased by ~30% when the Ile89Leu variant was transfected (Figure 3). Analysis of the 3D model of PLAP based on its 3D structure shows that the residue Ile89 is on the surface of the molecule and close to the homodimer interface (Figure 4).

DISCUSSION

A genetic analysis of a mouse model of interspecific congenic strains, comforted results from the literature (Dehghani et al., 2000) indicating that *Alppl2* is an important gene for early development. This prompted us to study this gene in human cases of RSA. In humans, *Alppl2* has two orthologs: *ALPP* and *ALPPL2* encoding placental (PLAP) and germinal (GCAP) alkaline phosphatases, respectively. PLAP is synthesized in the placenta as early as the seventh gestational week. PLAP is a dimeric glycosyl-phosphatidylinositol-anchored (GPI) protein that catalyses the hydrolysis of phosphomonoesters. Its natural substrates are largely unknown although it has been shown to also dephosphorylate bone alkaline phosphatase substrates phosphoethanolamine, inorganic pyrophosphate and pyridoxal 5'-phosphate (Whyte et al., 1995). Its physiological role remains elusive. Studies performed in different models have shown that it could be involved in cell proliferation during pregnancy (She et al., 2000a; She et al., 2000b). In 2011, Bellazi and coworkers confirmed that PLAP exerts positive proliferative effects on trophoblastic cells (Bellazi et al., 2010).

In the present study, we investigated the association between *ALPP* and the occurrence of idiopathic recurrent pregnancy losses in Mediterranean women, ethnically homogeneous. We found six SNPs as significantly associated with miscarriage (Table 2). Two of them were located in the coding region: Ile89Leu (rs130226692), which is a negative risk factor (the protecting effect culminating in homozygous women) and Tyr268Tyr (rs2260309) which is a positive risk factor (Table 4). Ten haplotypes were associated with recurrent pregnancy losses: five acting as negative and five as positive risk factors (Table 6). Among the protecting haplotypes, "CACGGCACCCCTC**T**CCTAC" (Block 2) is the most significant

with a p value reaching 0.0017. This result is correlated with SNPs analysis since it contains the protecting SNP Ile89Leu (in bold, rs130226692). Another haplotype, "TCA**G**CC", which favors the risk of miscarriages, contains the SNP Tyr268Tyr (in bold, rs2260309). Concerning crosses between the different haplotypes, two combinations (B and Q; Table 9) are independently associated with recurrent miscarriages, the Q combination being associated with a five-fold mean increase in the clinical risk, rendering it today one of the best predictive markers for the risk of recurrent pregnancy loss.

As expected for a polymorphism the position Ile89 is not conserved in various alkaline phosphatases from various species and no mutation was reported at this position in human TNAP responsible for hypophosphatasia. Analysis of the 3D model of PLAP based on its 3D structure shows that the residue Ile89 is on the surface of the molecule and close to the homodimer interface (Figure 4). It is noticeable that residues in the same area have been shown crucial to prevent non tissue-specific/tissue-specific alkaline phosphatases dimerization and to allow dimerization between paralog isoforms (Le Du and Millan, 2002). Since the catalytic activity of alkaline phosphatases depends on a dimeric configuration, an effect of the Ile89Leu substitution on the dimer stability, for instance, cannot be ruled out. However, in the present study we showed that the variant Ile89Leu expressed more alkaline phosphatase activity than the reference variant Pl^1 and further kinetic studies will be needed to clarify the effect of this variant at the biochemical level The putative effect of the variation rs2260309 is still more unclear. *In silico*, however, the SNP is predicted by the web server ESE Finder (http://rulai.cshl.edu/tools/ESE) to decrease the score of an exonic splicing enhancer SRp55 from 3.8 (allele C) to 2.8 (allele T), suggesting that the gene carrying allele C could more easily undergo correct splicing than allele T. This hypothesis will deserve further analysis on the proportion of spliced mRNA variants characteristic of each allele.

These haplotypes, grouped together, are made up of many SNPs located in introns which may destabilize the pre-mRNA and therefore the transcription. However further studies, including analysis of the promoter, will be needed to confirm this. We have also previously seen, in the univariable analysis, that four intronic SNPs were associated with miscarriages. These variant alleles, as well as other non-significant intronic SNPs (Table 2) are not described in the Pubmed database (http://www.ncbi.nlm.nih.gov/SNP/) and may therefore be specific of our Mediterranean population.

Furthermore, as the placenta is made up of both maternal and paternal genomes, it would be interesting to study these SNPs and haplotypes among fathers in order to assess the links between placental genotypes and pregnancy loss.

As mentioned above, PLAP is a GPI–anchored protein and these proteins serve a variety of functions that include adhesion, receptors, signal transduction and complement activation. Lange and co-workers described three cases of women with paroxysmal nocturnal hemoglobinuria (PNH), a rare haematological disorder. One patient had two spontaneous miscarriages, and a 30% rate of spontaneous abortions or stillbirths was previously described in PNH women (Langer et al., 1998). Although PNH is an acquired pathology in the X-linked *PIGA* gene, it would be interesting to study the occurrence of miscarriages in women with PNH, especially women expecting a boy (hemizygous placenta). Moreover, rare cases of inherited GPI deficiency have also been described (Almeida et al., 2009; Krawitz et al., 2010). Mutations in PIGV, another gene involved in GPI anchoring, have been shown to be responsible for Marbry Disease, an inherited recessive disorder characterized by a high level of serum bone/liver alkaline phosphatase (Krawitz et al., 2010). Interestingly, the original publication of Marbry reported three cases of miscarriage in the affected pedigree (Mabry et al., 1970). All these results support the hypothesis of a role of PLAP in RSA.

Previous studies conducted in Northern Europe (Finland and Sweden) have already associated *ALPP* polymorphisms with miscarriage. Beckman, as early as 1972, showed that two polymorphisms, P209R and E429G, modify the biochemical properties of the enzyme (Beckman et al., 1972; Wennberg et al., 2002). In the present study, however, these SNPs were not found, maybe due to the geographical origin of our population (Mediterranean, Southern Europe). In addition, on the wild- type allele, we find an alanine instead of a proline for the first SNP; and a proline instead of a glutamic acid as in the reference sequence in Pubmed (http://www.ncbi.nlm.nih.gov). These data confirm the crucial importance of the patients' geographical origin included in linkage analysis attempts between common clinical disorders and genetic characteristics. Similarly, in 2010, Bellazi and coworkers referred to linkage disequilibrium between a sequence variation (rs2014683) in the promoter and the P209R polymorphism (Bellazi et al., 2010).

CONCLUSION

The genetic of complex diseases is a very challenging issue, since the phenotype is the result of complex interactions between numerous proteins, encoded by various genes. Therefore it is important to diversify the approaches in order to discover the different actors. Herein, the analysis of an unusual mouse model provides trails that directly brought us to a promising candidate. The variants identified here in ALPP may account for ~10% of the cases of recurrent spontaneous abortion in the studied population. With the progress in whole genome sequencing, the setting up of a database of ‘at-risk’ variants for complex disease is the ground for a novel personalized medicine. We hope that the present findings may contribute to this aim in the future.

Acknowlegments

The authors thank Drs E. Nouvellon and E. Mercier for their skillful assistance and Ms E. Saurel for her help in sequence analysis.

REFERENCES

1. **Matzuk MM, Lamb DJ 2008 The biology of infertility: research advances and clinical challenges. Nat Med 14:1197-1213**
2. **Rai R, Regan L 2006 Recurrent miscarriage. Lancet 368:601-611**
3. **Hogge WA, Byrnes AL, Lanasa MC, Surti U 2003 The clinical use of karyotyping spontaneous abortions. Am J Obstet Gynecol 189:397-400; discussion 400-392**
4. **Jauniaux E, Farquharson RG, Christiansen OB, Exalto N 2006 Evidence-based guidelines for the investigation and medical treatment of recurrent miscarriage. Hum Reprod 21:2216-2222**
5. **Kupferminc MJ 2003 Thrombophilia and pregnancy. Reprod Biol Endocrinol 1:111**
6. **L'Hote D, Serres C, Laissue P, Oulmouden A, Rogel-Gaillard C, Montagutelli X, Vaiman D 2007 Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. Genetics 176:1907-1921**
7. **Burgio G, Szatanik M, Guenet JL, Arnau MR, Panthier JJ, Montagutelli X 2007 Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics 177:2321-2333**
8. **L'Hote D, Laissue P, Serres C, Montagutelli X, Veitia RA, Vaiman D 2010 Interspecific resources: a major tool for quantitative trait locus cloning and speciation research. Bioessays 32:132-142**
9. **L'Hôte D, Vatin M, Auer J, Castille J, Passet B, Montagutelli X, Serres C, Vaiman D 2011 Fidgetin-Like1 Is a Strong Candidate for a Dynamic Impairment of Male Meiosis Leading to Reduced Testis Weight in Mice. PLoS One:in press**
10. **Laissue P, Burgio G, l'Hote D, Renault G, Marchiol-Fournigault C, Fradelizi D, Fellous M, Serres C, Montagutelli X, Monget P, Vaiman D 2009 Identification of Quantitative Trait Loci responsible for embryonic lethality in mice assessed by ultrasonography. Int J Dev Biol 53:623-629**
11. **Dehghani H, Narisawa S, Millan JL, Hahnel AC 2000 Effects of disruption of the embryonic alkaline phosphatase gene on preimplantation development of the mouse. Dev Dyn 217:440-448**
12. **Wennberg C, Kivela A, Holmgren PA 1995 Placental and germ cell alkaline phosphatase RFLPs and haplotypes associated with spontaneous abortion. Hum Hered 45:272-277**

13. **Wennberg C, Kozlenkov A, Di Mauro S, Frohlander N, Beckman L, Hoylaerts MF, Millan JL 2002 Structure, genomic DNA typing, and kinetic characterization of the D allozyme of placental alkaline phosphatase (PLAP/ALPP). Hum Mutat 19:258-267**
14. **Bellazi L, Germond S, Dupont C, Brun-Heath I, Taillandier A, De Mazancourt P, Dieudonne MN, Mornet E 2010 A sequence variation in the promoter of the placental alkaline phosphatase gene (ALPP) is associated with allele-specific expression in human term placenta. Placenta 31:764-769**
15. **Ferianec V, Linhartova L 2011 Extreme elevation of placental alkaline phosphatase as a marker of preterm delivery, placental insufficiency and low birth weight. Neuro Endocrinol Lett 32:154-157**
16. **Whyte MP, Landt M, Ryan LM, Mulivor RA, Henthorn PS, Fedde KN, Mahuren JD, Coburn SP 1995 Alkaline phosphatase: placental and tissue-nonspecific isoenzymes hydrolyze phosphoethanolamine, inorganic pyrophosphate, and pyridoxal 5'-phosphate. Substrate accumulation in carriers of hypophosphatasia corrects during pregnancy. J Clin Invest 95:1440-1445**
17. **Fauvert D, Brun-Heath I, Lia-Baldini AS, Bellazi L, Taillandier A, Serre JL, de Mazancourt P, Mornet E 2009 Mild forms of hypophosphatasia mostly result from dominant negative effect of severe alleles or from compound heterozygosity for severe and moderate alleles. BMC Med Genet 10:51**
18. **Le Du MH, Stigbrand T, Taussig MJ, Menez A, Stura EA 2001 Crystal structure of alkaline phosphatase from human placenta at 1.8 A resolution. Implication for a substrate specificity. J Biol Chem 276:9158-9165**
19. **She QB, Mukherjee JJ, Chung T, Kiss Z 2000 Placental alkaline phosphatase, insulin, and adenine nucleotides or adenosine synergistically promote long-term survival of serum-starved mouse embryo and human fetus fibroblasts. Cell Signal 12:659-665**
20. **She QB, Mukherjee JJ, Huang JS, Crilly KS, Kiss Z 2000 Growth factor-like effects of placental alkaline phosphatase in human fetus and mouse embryo fibroblasts. FEBS Lett 469:163-167**
21. **Le Du MH, Millan JL 2002 Structural evidence of functional divergence in human alkaline phosphatases. J Biol Chem 277:49808-49814**
22. **Langer B, Simeoni U, Schlaeder G 1998 Prognostic criteria for fetal pyelectasis. Ultrasound Obstet Gynecol 11:82-83**

23. Almeida A, Layton M, Karadimitris A 2009 Inherited glycosylphosphatidyl inositol deficiency: a treatable CDG. Biochim Biophys Acta 1792:874-880

24. Krawitz PM, Schweiger MR, Rodelsperger C, Marcelis C, Kolsch U, Meisel C, Stephani F, Kinoshita T, Murakami Y, Bauer S, Isau M, Fischer A, Dahl A, Kerick M, Hecht J, Kohler S, Jager M, Grunhagen J, de Condor BJ, Doelken S, Brunner HG, Meinecke P, Passarge E, Thompson MD, Cole DE, Horn D, Roscioli T, Mundlos S, Robinson PN 2010 Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. Nat Genet 42:827-829

25. Mabry CC, Bautista A, Kirk RF, Dubilier LD, Braunstein H, Koepke JA 1970 Familial hyperphosphatase with mental retardation, seizures, and neurologic deficits. J Pediatr 77:74-85

26. Beckman L, Beckman G, Magnusson SS 1972 Relationship between placental alkaline phosphatase phenotypes and the frequency of spontaneous abortion in previous pregnancies. Hum Hered 22:15-17

FIGURE LEGENDS

Figure 1: Haploview LD Plot®. The 49 SNPs are represented in the abscissa (1 to 49). Each box shows the strength of association. The boxes colored in red indicate the haplotypes associated with each other. Several haplotypes appeared to be related. In block 2, bond strength is important because of the proximity of SNPs (SNP7: *Alpp* 988, SNP8: *Alpp* 989, SNP9: *Alpp* 991, SNP10: *Alpp* 992, SNP11: *Alpp* 994).

Figure 2: Cross between different haplotypes (Haploview®). Each haplotype is associated with its frequency. The numbers (1 to 49) correspond to the different SNPs.

Figure 3: *In vitro* alkaline phosphatase activity of Pl^l and Ile89Leu variants of ALPP gene after transfection in COS-7 cells. Alkaline phosphatase activity was normalized with β-galactosidase activity as described in Materials and Methods. COS-7 cells transfected with empty pcDNA3.1 were used as negative control of alkaline phosphatase activity. In each case, experiments were repeated six times independently. Bars mean ± SEM of the data. ** $p<0.01$, paired t-test**Figure 4:** 3D model of the human PLAP monomer. The residue Ile89 is shown on the surface of the molecule (yellow). Blue and red residues correspond to basic and acidic residues, respectively .

Table 1: Characteristics of the population analyzed

	Cases (n=100)	Controls (n=100)
Average âge (years)	31,72	51,47
Average number of miscarriages (per women)	2,51	0 (reported)
Average number of children (per women)	0	2,6
Early micarriages (< 10 weeks of gestation)	33	-
Late miscarriages (≥ 10 weeks of gestation)	54	-
Early and late miscarriages	13	-

Table 2: SNPs and results of univariable analysis. For each SNP, heterozygous and homozygous genotypes are grouped. OR: Odds Ratio; 95% CI: Confidence Interval; * Significant results; *Alpp*: SNP position on the gene.

	Alpp	SNP	Intron/Protein	Cases	Controls	p (logistic regression)	OR	95% CI	p (Chi Square)
1	691	C/T	Before ATG	0	1		1	[0,062;16,217]	>0,9999
	691	T/T	Before ATG	1	0				
2	805	T/C	Before ATG	1	0	0,9782	201377,628		>0,9999
3	943	C/T (rs1130335)	Pro25Leu	21	8	0,09	1,949	[0,901;4,217]	0,1275
	943	T/T (rs1130335)	Pro25Leu	0	4				
4	967	C/T	Intron	1	0	0,9782	201377,628		>0,9999
5	977	A/C	Intron	0	4	0,9713	1,77E-06		0,1297
6	978	C/A	Intron	0	4	0,9713	1,77E-06		0,1297
7	988	G/C	Intron	19	8	0,0268 *	2,698	[1,121;6,493]	0,0385
8	989	G/A	Intron	19	8	0,0268 *	2,698	[1,121;6,493]	0,0385
9	991	C/G	Intron	4	3	0,7014	1,347	[0,294;6,181]	>0,9999
10	992	A/G	Intron	4	1	0,2087	4,125	[0,453;37,581]	0,365
11	994	C/A	Intron	4	1	0,2087	4,125	[0,453;37,581]	0,365
12	1002	A/A (rs3748969)	Intron	0	2	0,9797	1,81E-06		0,4773
13	1017	T/C (rs2014448)	Intron	4	7	0,0607	0,515	[0,257;1,030]	0,0852
	1017	C/C (rs2014448)	Intron	69	77				
14	1029	C/A	Intron	1	0	0,9782	201377,628		>0,9999
15	1153	T/C (rs1130338)	Asp64Asp	5	10	0,1877	0,474	[0,156;1,440]	0,2829
16	1264	C/T	Intron	1	1		1	[0,062;16,217]	>0,9999
17	1338	A/T (rs13026692)	Ile89Leu	15	19	0,0162 *	0,474	[0,258;0,871]	0,023
	1338	T/T (rs13026692)	Ile89Leu	9	21				
18	1395	C/G (rs142439881)	Intron	2	0	0,9797	552989,648		0,4773
19	1434	C/A	Intron	1	0	0,9782	201377,628		>0,9999
20	1447	T/C	Intron	1	0	0,9782	201377,628		>0,9999
21	1484	A/G	Intron	2	0	0,9797	552989,648		0,4773
22	1874	C/A (rs150558405)	Ala176Ala	1	2	0,5684	0,495	[0,044;5,549]	>0,9999
23	1943	C/G (rs2853374)	Ser199Ser	6	7	0,302	0,574	[0,200;1,646]	0,4343
	1943	G/G (rs2853374)	Ser199Ser	84	87				
24	1973	C/C (rs17412770)	Ala209Ala	1	0	0,9782	201377,628		>0,9999
25	2134	T/C (rs12620832)	Intron	17	2	0,0825	2,023	[0,913;4,481]	0,118
	2134	C/C (rs12620832)	Intron	3	9				
26	2200	G/A	Intron	6	2	0,3154	2,064	[0,501;8,494]	0,4951
	2200	A/A	Intron	0	1				
27	2287	G/C (rs1048988)	Arg231Pro	30	21	0,1681	1,533	[0,835;2,812]	0,2192
	2287	C/C (rs1048988)	Arg231Pro	5	5				
28	2383	G/A (rs2853378)	Arg263His	0	2	0,9797	1,81E-06		0,4773
29	2402	C/T	Intron	5	0	0,1352	5,211	[0,598;45,436]	0,2137
	2402	T/T	Intron	0	1				
30	2447	A/G	Intron	26	22	0,671	1,128	[0,647;1,965]	0,777
	2447	G/G	Intron	23	24				
31	2466	T/C	Intron	1	0	0,9782	201377,628		>0,9999
32	2498	T/C (rs2260309)	Tyr268Tyr	31	16	0,0049 *	2,253	[1,279;3,970]	0,0072
	2498	C/C (rs2260309)	Tyr268Tyr	31	26				
33	2585	A/G	Intron	6	0	0,977	1567149,781		0,0382
34	2587	C/G	Intron	4	0	0,9713	564510,265		0,1297
35	2591	C/G	Intron	0	1	0,9782	4,97E-06		>0,9999
36	2599	G/G	Intron	2	0	0,9797	552989,648		0,4773
37	2621	T/C	Intron	14	5	0,0372 *	3,093	[1,069;8,946}	0,0537
38	2623	T/C	Intron	13	8	0,253	1,718	[0,679;4,348]	0,3562
39	2671	A/T	Intron	4	0	0,9713	564510,265		0,1297
40	3175	C/G	Intron	1	0	0,9782	201377,628		>0,9999
41	3217	A/C	Intron	4	0	0,9713	564510,265		0,1297
42	3315	C/A	Intron	18	26	0,0472 *	0,538		0,0653
	3315	A/A	Intron	6	11			[0,291;0,992]	
43	3463	G/T	Intron	2	0	0,9797	552989,648		0,4773
44	3510	C/T	Intron	1	0	0,9782	201377,628		>0,9999
45	3539	T/C	Intron	10	8	0,8095	1,123	[0,436;2,895]	>0,9999
	3539	C/C (rs73001984)	Intron	0	1				
46	3578	C/T (rs142493383)	Arg442Tryp	0	2	0,9797	1,81E-06		0,4773
47	3756	C/G (rs2981374)	Pro501Arg	2	1	0,5684	2,02	[0,180;22,651]	>0,9999
48	3789	G/C	Arg512Pro	2	0	0,9797	552989,648		0,4773
49	3823	C/A	Ala523Ala	2	0	0,9797	552989,648		0,4773

Table 3: Distribution of significant SNPs according to recurrent miscarriages clinical subtypes.

Alpp	**Number of women with miscarriages <10 weeks of gestation (n=33)**	**Number of women with miscarriages >10 weeks of gestation (n=54)**	**Number of women with both miscarriages (n=13)**
988	5 (15%)	9 (17%)	5 (38%)
989	5 (15%)	9 (17%)	5 (38%)
1338	7 (21%)	14 (26%)	3 (23%)
2498	18 (55%)	34 (63%)	10 (77%)
2621	4 (12%)	10 (19%)	0 (0%)
3315	8 (24%)	15 (28%)	1 (8%)

Table 4: Multivariable analysis. * Significant results. These two SNPs were independently associated with miscarriage. The first one, which is a synonymous SNP (*Alpp* 2498, Tyr268Tyr, rs2260309) is a positive risk factor and the second one, a missense NP, (*Alpp* 1338, Ile89Leu, rs130226692) is protective.

Alpp	**SNP**	**p (logistic regression)**	**OR**	**95% CI**
2498	Tyr268Tyr (rs2260309)	0,0027 *	2,824	[1,432;5,571]
1338	Ile89Leu (rs13026692)	0,0004 *	0,277	[0,136;0,564]

Table 5: Influence of genotypes for the two significant polymorphisms in multivariable analysis (rs2260309 and rs130226692). * Significant results.

SNP	Cases	Controls	p (logistic regression)	OR	95% CI
Tyr268Tyr (rs2260309) T/C	31	16	0,0026 *	3,059	[1,478;6,332]
Ile89Leu (rs13026692) A/T	31	26	0,0371 *	2,039	[1,043;3,986]
Tyr268Tyr (rs2260309) C/C	15	19	0,2208	0,623	[0,292;1,329]
Ile89Leu (rs13026692) T/T	9	21	0,0126 *	0,338	[0,144;0,793]

Table 6: Statistical analysis of the different haplotypes (Haploview®). * Significant result

Haplotype	Frequency	Case, Control Ratio Counts	Case,Control Frequencies	Chi Square	P Value	Protective/Risk factor
Block 1						
CTC	0.897	176.0 : 24.0, 183.0 : 17.0	0.880, 0.915	1.332	0.2485	
CTT	0.093	21.0 : 179.0, 16.0 : 184.0	0.105, 0.080	0.745	0.3882	
Block 2						
CACGGCACCCCTCACCTAC	0.464	96.0 : 104.0, 89.7 : 110.3	0.480, 0.448	0.401	0.5267	
CACGGCACCCCTCTCCTAC	0.215	30.0 : 170.0, 55.8 : 144.2	0.150, 0.279	9.879	0.0017 *	Protective
CACGGCACCTCTCACCTAC	0.158	40.0 : 160.0, 23.3 : 176.7	0.200, 0.117	5.224	0.0223 *	Risk factor
CACCACACCTCTCACCTAC	0.045	13.0 : 187.0, 5.0 : 195.0	0.065, 0.025	3.723	0.0537	
CACGGCACCCCCCACCTAC	0.034	5.0 : 195.0, 8.5 : 191.5	0.025, 0.043	0.945	0.3311	
CACCAGGACTCTCACCTAC	0.013	4.0 : 196.0, 1.0 : 199.0	0.020, 0.005	1.823	0.177	
CCAGGCACCTCTCACCTAC	0.010	0.0 : 200.0, 4.0 : 196.0	0.000, 0.020	4.04	0.0444 *	Protective
Block 3						
GTTGGGC	0.611	111.3 : 88.7, 133.0 : 67.0	0.557, 0.665	4.935	0.0263 *	Protective
GTTGCGC	0.166	36.2 : 163.8, 30.0 : 170.0	0.181, 0.150	0.702	0.402	
CTTGGGC	0.099	23.5 : 176.5, 16.0 : 184.0	0.118, 0.080	1.583	0.2083	
GTCGGGC	0.073	13.1 : 186.9, 16.0 : 184.0	0.066, 0.080	0.305	0.5808	
GTCAGGC	0.013	5.0 : 195.0, 0.0 : 200.0	0.025, 0.000	5.087	0.0241 *	Risk factor
Block 4						
TTACCC	0.580	101.0 : 99.0, 131.0 : 69.0	0.505, 0.655	9.236	0.0024 *	Protective
TCACCC	0.380	84.0 : 116.0, 68.0 : 132.0	0.420, 0.340	2.716	0.0993	
TTGCCC	0.010	4.0 : 196.0, 0.0 : 200.0	0.020, 0.000	4.04	0.0444 *	Risk factor
TCAGCC	0.010	4.0 : 196.0, 0.0 : 200.0	0.020, 0.000	4.04	0.0444 *	Risk factor
Block 5						
ACAC	0.785	162.0 : 38.0, 152.0 : 48.0	0.810, 0.760	1.481	0.2236	
ACAA	0.195	30.0 : 170.0, 48.0 : 152.0	0.150, 0.240	5.16	0.0231 *	Protective
TCAC	0.010	4.0 : 196.0, 0.0 : 200.0	0.020, 0.000	4.04	0.0444 *	Risk factor
Block 6						
GCTCCGC	0.935	187.0 : 13.0, 187.0 : 13.0	0.935, 0.935	0.0	1.0	
GCCCCGC	0.045	8.0 : 192.0, 10.0 : 190.0	0.040, 0.050	0.233	0.6295	

Table 7: Representation of the 18 combinations with a frequency between 10.83% and 1%.

* Significant results ° Results near the significance threshold.

Combinations	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Frequency of Haplotype	p Chi Square
A	CTC	CACGGCACCCCTCACCTAC	GTTGGGC	TTACCC	ACAC	GCTCCGC	10,83%	0,8774
B	CTC	CACGGCACCCCTCACCTAC	GTTGGGC	TCACCC	ACAC	GCTCCGC	7,09%	0,0646 °
C	CTC	CACGGCACCCCTCTCCTAC	GTTGGGC	TTACCC	ACAC	GCTCCGC	5,01%	0,0646 °
D	CTC	CACGGCACCTCTCACCTAC	GTTGGGC	TTACCC	ACAC	GCTCCGC	3,69%	> 0,9999
E	CTC	CACGGCACCCCTCTCCTAC	GTTGGGC	TCACCC	ACAC	GCTCCGC	3,23%	0,541
F	CTC	CACGGCACCCCTCACCTAC	GTTGCGC	TTACCC	ACAC	GCTCCGC	2,94%	0,8471
G	CTC	CACGGCACCCCTCACCTAC	GTTGGGC	TTACCC	ACAA	GCTCCGC	2,69%	0,4218
H	CTC	CACGGCACCTCTCACCTAC	GTTGGGC	TCACCC	ACAC	GCTCCGC	2,41%	0,3562
I	CTC	CACGGCACCCCTCACCTAC	GTTGCGC	TCACCC	ACAC	GCTCCGC	1,92%	0,3082
J	CTC	CACGGCACCCCTCACCTAC	GTTGGGC	TCACCC	ACAA	GCTCCGC	1,76%	0,8277
K	CTC	CACGGCACCCCTCACCTAC	CTTGGGC	TTACCC	ACAC	GCTCCGC	1,75%	0,4951
L	CTC	CACGGCACCCCTCTCCTAC	GTTGCGC	TTACCC	ACAC	GCTCCGC	1,36%	0,4951
M	CTC	CACGGCACCCCTCACCTAC	GTCGGGC	TTACCC	ACAC	GCTCCGC	1,29%	0,9999
N	CTC	CACGGCACCCCTCTCCTAC	GTTGGGC	TTACCC	ACAA	GCTCCGC	1,25%	0,0407 *
O	CTC	CACGGCACCCCTCACCTAC	CTTGGGC	TCACCC	ACAC	GCTCCGC	1,14%	0,2137
P	CTT	CACGGCACCCCTCACCTAC	GTTGGGC	TTACCC	ACAC	GCTCCGC	1,12%	> 0,9999
Q	CTC	CACCACACCTCTCACCTAC	GTTGGGC	TTACCC	ACAC	GCTCCGC	1,05%	0,019 *
R	CTC	CACGGCACCTCTCACCTAC	GTTGCGC	TTACCC	ACAC	GCTCCGC	1,00%	> 0,9999

Table 8: Univariable analysis on the aforementioned four haplotype combinations. All of these four combinations are significant. Two combinations are protective and two are positive risk factors for miscarriages.

Combinations	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	p (logistic regression)	OR	95% CI
B	CTC	CACGGCAC CCCTCACCTAC	GTTG GGC	TCA CCC	ACAC	GCTC CGC	0,0458 *	1,994	[1,013; 3,926]
C	CTC	CACGGCAC CCCTCTCCTAC	GTTG GGC	TTA CCC	ACAC	GCTC CGC	0,0458 *	0,501	[0,255; 0,987]
N	CTC	CACGGCAC CCCTCTCCTAC	GTTG GGC	TTA CCC	ACAA	GCTC CGC	0,0443 *	0,116	[0,014; 0,947]
Q	CTC	CACCACAC CTCTCACCTAC	GTTG GGC	TTA CCC	ACAC	GCTC CGC	0,0166 *	4,831	[1,332; 17,524]

Table 9: Results of the stepwise multivariable analysis performed on the aforementioned four haplotype combinations. * Significant results. Two combinations ("B" and "Q") are independent positive risk factors for miscarriages.

Combinations	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	p (logistic regression)	OR	95% CI
B	CTC	CACGGCACC CCTCACCTAC	GTTG GGC	TCA CCC	ACAC	GCTC CGC	0,0174 *	2,346	[1,162; 4,737]
C	CTC	CACGGCAC CCCTCTCCTAC	GTTG GGC	TTA CCC	ACAC	GCTC CGC	0,4485	0,753	[0,362; 1,567]
N	CTC	CACGGCAC CCCTCTCCTAC	GTTG GGC	TTA CCC	ACAA	GCTC CGC	0,0897	0,152	[0,017; 1,339]
Q	CTC	CACCACAC CTCTCACCTAC	GTTG GGC	TTA CCC	ACAC	GCTC CGC	0,0137 *	5,207	[1,401; 19,349]

Figure 1:

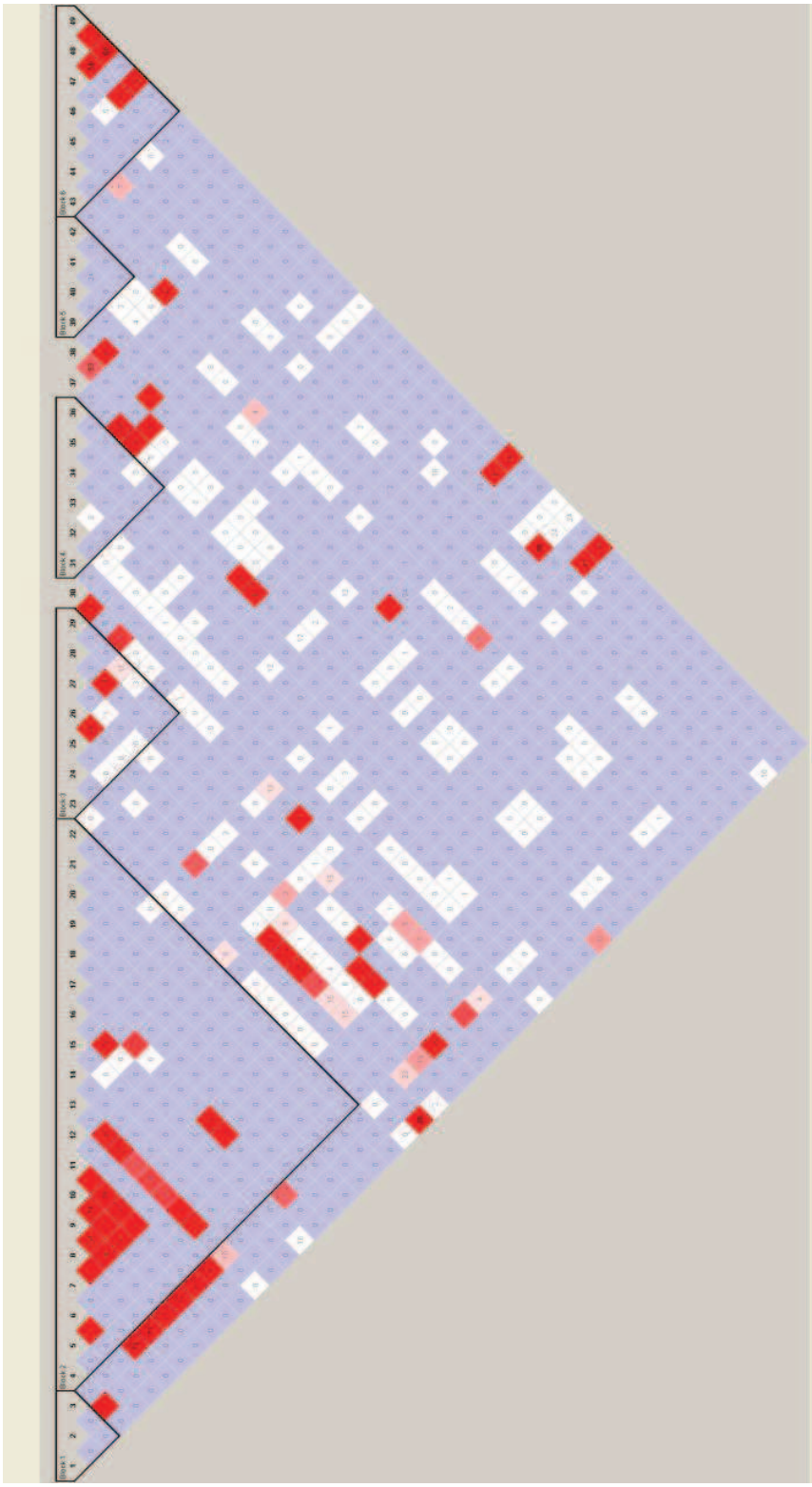


Figure 2:

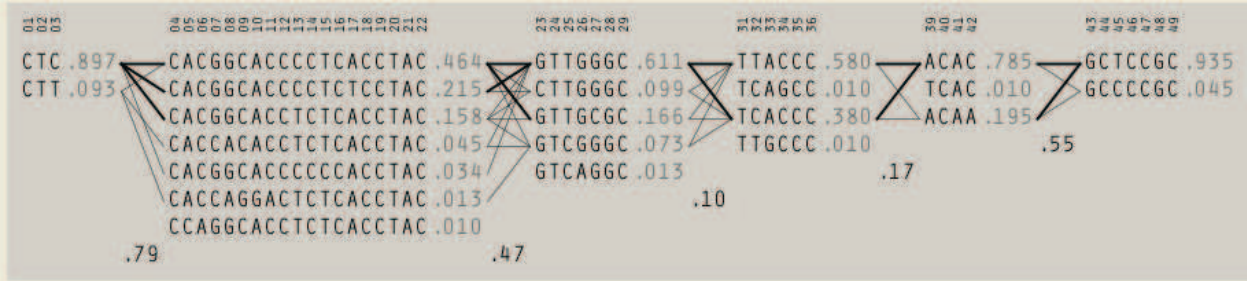

CTC .897
CTT .093
CACGGCACCCCTCACCTAC .464
CACGGCACCCCTCTCCTAC .215
CACGGCACCTCTCACCTAC .158
CACCACACCTCTCACCTAC .045
CACGGCACCCCCCACCTAC .034
CACCAGGACTCTCACCTAC .013
CCAGGCACCTCTCACCTAC .010
.79
GTTGGGC .611
CTTGGGC .099
GTTGCGC .166
GTCGGGC .073
GTCAGGC .013
.47
TTACCC .580
TCAGCC .010
TCACCC .380
TTGCCC .010
.10
ACAC .785
TCAC .010
ACAA .195
.17
GCTCCGC .935
GCCCCGC .045
.55

Figure 3: X

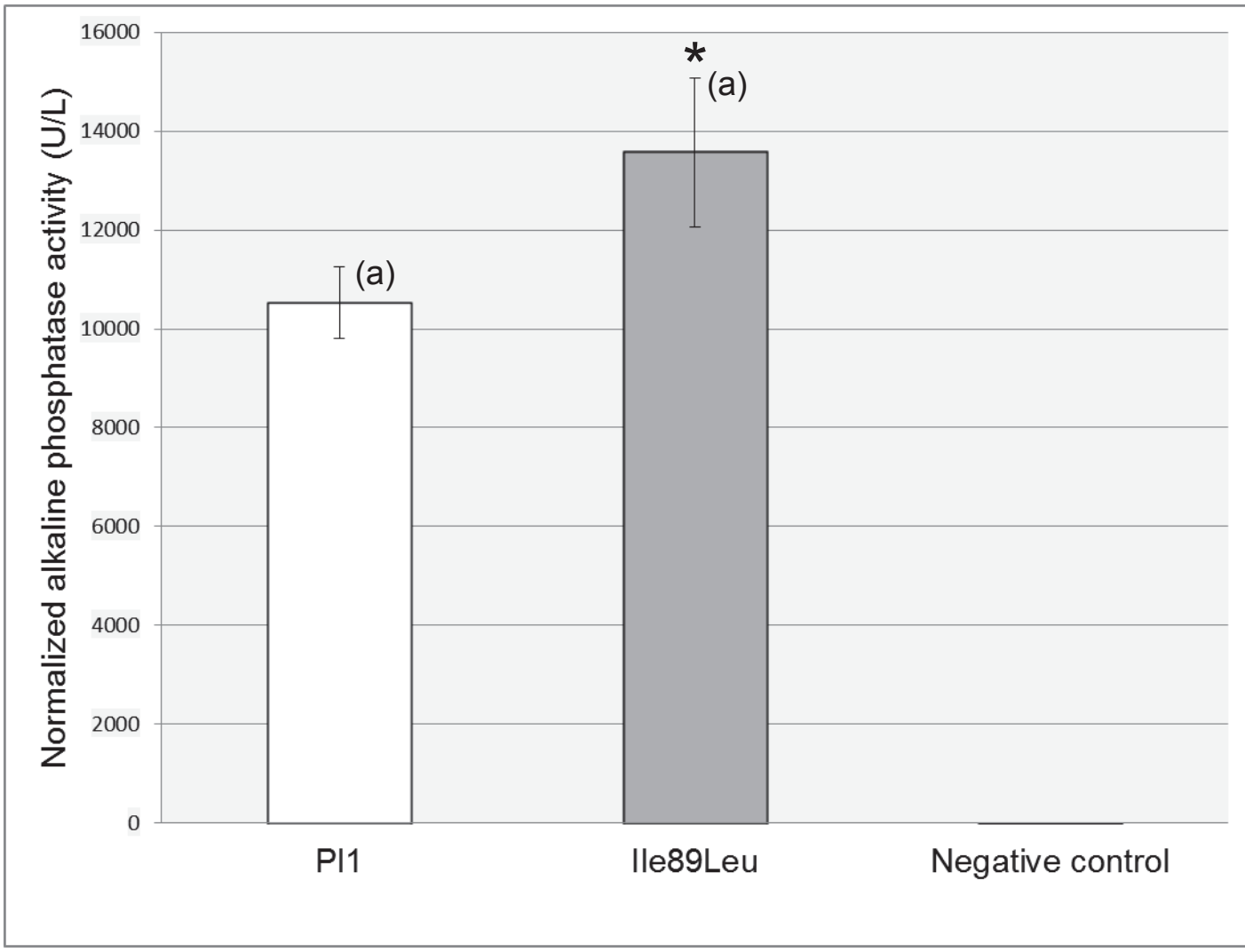
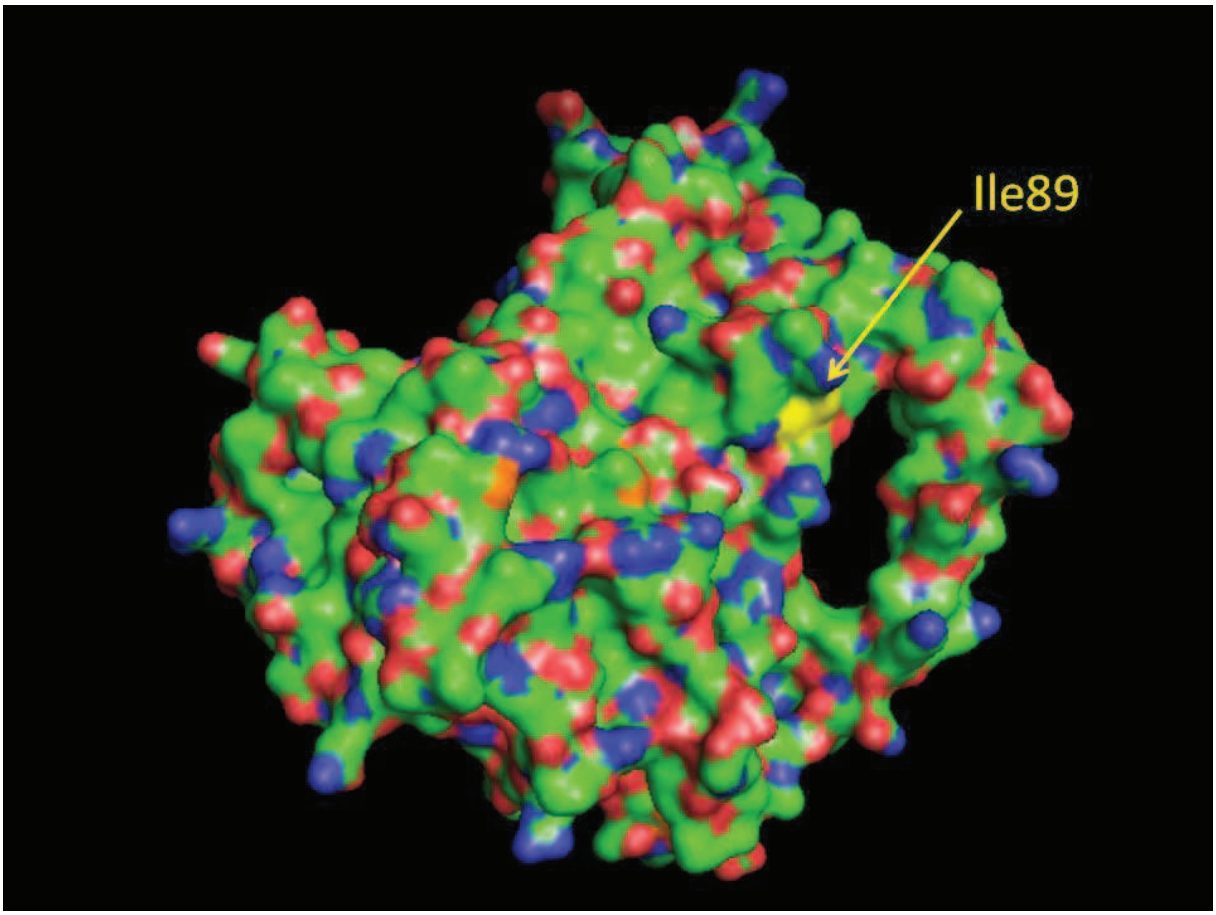

Normalized alkaline phosphatase activity (U/L)
16000
14000
12000
10000
8000
6000
4000
2000
0
(a)
*
(a)
PI1
Ile89Leu
Negative control

Figure 4 :



Article 5

Identification of new QTL region on mouse chromosome 1 responsible for a partial globozoospermia

(Résultats préliminaires)

Résultats préliminaires

Dans le cadre de travaux récents, j'ai réalisé le phénotypage des lignées recombinantes obtenues à partir de 66HMMU1 par échographie à haute fréquence et j'ai observé une diminution significative de la moyenne du nombre d'embryons implantés dans le croisement homozygote de la lignée IRCS Rc3 comparée au croisement contrôle homozygote B6. J'ai démontré que le phénotype observé résulte d'un défaut de fertilité des mâles Rc3, à savoir une globozoospermie partielle.

En effet, par la réalisation de FIV en utilisant les spermatozoïdes Rc3, j'ai observé une diminution significative du nombre d'ovocytes fécondés. De plus, l'observation de la morphologie des spermatozoïdes a révélé une anomalie au niveau de la tête à savoir une réduction de la taille de l'acrosome voire son absence sur une partie de la population spermatique. Les premières observations en microscopie électronique semblent confirmer ces résultats relatifs au défaut d'acrosome mais révèlent également la présence en grand nombre de gouttelettes lipidiques au niveau du flagelle.

Dans le but d'identifier la région génomique contenant le gène causal, j'ai réalisé différents croisements homozygotes et la mise en relation génotype/phénotype a permis de réduire la région QTL à une région de 6Mb sur le chromosome 1. Dans le but de proposer des gènes candidats, la même approche que celle des précédentes études est employée en se basant sur leur niveau d'expression testiculaire, sur la présence de SNP entre SEG et B6 et sur les données de la littérature. Ces travaux sont actuellement en cours de réalisation.

Identification of new QTL region on mouse chromosome 1 responsible for a partial globozoospermia

Magalie Vatin, Virginie Firlej, Carmen Marchiol, Come Ialy-radio, Jana Auer, Daniel Vaiman, Catherine Serres and Ahmed Ziyyat

Université Paris Descartes, Institut Cochin Inserm U1016 CNRS UMR 8104, Paris, France.

Abstract

Globozoospermia is rare but severe form of teratozoospermia meanly characterized by round-head spermatozoa that lack an acrosome. This disorder in male infertility takes his origin in a disturbed spermatogenesis, which is expected to be induced by genetic factors. To date, mutation in two genes including *Spata16* have been identified as responsible for globozoospermia. The phenotyping of the Interspecific Recombinant Congenic strains revealed a diminution of sperm fertilizing ability due to a severe teratozoospermia. The similarity in the human pathology of globozoospermia and the mouse phenotype permitted us to propose three candidate genes specifically expressed in spermatocytes and spermatids: *dnajb3*, *efhd1* and *spata3*.

Introduction

In approximately half of the 15% of couples who cannot conceive, the cause is ascribed to male infertility, a multifactorial syndrome encompassing a wide variety of syndrome. Causes of male infertility may be classified into four categories: defects in sperm production, reproductive tract obstruction, inflammation and sexual disorders. In more than half infertile men, the cause is idiopathic and may be due to genetic factors. Genetic defects leading to male infertility often affect spermatogenesis or sperm function .

Spermatogenesis is one of the most efficient cell-providing system in the adult mammalian organism (Sharpe RM, McKinnell C, 2003). This biological process is composed of a series of highly complex cellular events. It can be broadly divided into several discrete events: (1) renewal of spermatogonial stem cells and spermatogonia via mitosis, (2) proliferation (via mitosis) and differentiation of spermatogonia, (3) meiosis, (4) spermiogenesis (transformation of round spermatids into elongated spermatids and spermatozoa) and (5) spermiation (the release of sperm from the epithelium into the tubule lumen) (Cheng CY, Mruk DD.2010; O'Donnell L, Nicholls PK, 2011).

All the steps are highly regulated and dysfunctions in this key physiological process can result in infertility. Regulation of this complex process depends on the cooperation of many genes, which are expressed at the different stages . Problems during spermatogenesis are reflected in a lower or null production of spermatozoa and are described by routine semen analysis using terms such as azoospermia, oligozoospermia, teratozoospermia, asthenozoospermia or a combination of the last three (oligoasthenoteratozoospermia). Globozoospermia is a rare (incidence <0,1% in male infertile patients) form of teratozoospermia).

Identification of genes playing a crucial role in spermatogenesis is mainly based on observations in rodents, particularly in the mouse model. Mice appear to be an excellent model to study human infertility due to the conservation of the great majority of genes and processes involved in sperm production (Claire L. Borg, katja M Wolski, 2009). The production of haploid germ cells involves the coordinated expression of more than 2300 different genes (Schultz N, Hamra FK, 2003) and through knockout mouse models, at least 388 genes have been shown to be associated with spermatogenesis in mice (Massart A, Lissens W, 2012).

Teratozoospermia includes the deformation of the acrosome and the nucleus. During spermatogenesis diverse cellular and molecular processes allow sperm head formation and organization. Mutant models have improved the understanding of the etiology of teratozoospermia and clarified the mechanism involved in sperm head formation and organization (Venables JP, Cookes HJ, 2000; Escalier D 2001; Cookes HJ, Saunders PT, 2002). At the premeiotic and meiotic levels the genes from *Daz* family and *Tspy1* are highly expressed, however, during the postmeiotic stage the expression of the following genes, mainly responsible for sperm head formation and the normal sperm function, increased: *Prm2*, *Tnp1*, *Synj2*, and *Zpbp* (Fox MS, Ares VX, 2003).

In order to identify genetic causes of infertility, we used the interspecific recombinant congenic mouse strains (IRCS) developed at the Pasteur Institute (Paris, France). The IRCS

model harbors about 2% of *Mus spretus* (SEG/Pas) genome in a *Mus musculus domesticus* (C57BL6/J) background. The high polymorphism between *spretus* and *musculus* allows us to identify in the *spretus* region, QTL responsible for phenotypic variation in IRCS strain compared to the reference strain, C57BL6/J (B6) (Burgio G, Szatanik M,, 2007, L'Hôte D, Serres C, 2007, L'Hôte D, Laissue P, 2010). The phenotyping of the Rc3 strain permit us to identify spermatogenesis defects as acrosomal and flagellar anomalies absent in the referent strain. The combination of phenotypic and cartographic approaches allows us to rapidly map the QTL region on this unique *spretus* fragment on the chromosome 1 and to identify candidate genes.

Materials and methods

Ethics Statement

Procedures for handling and experimentation were conducted in accordance with the policies of the Paris Descartes University, the Cochin Institute and the Guidelines for Biomedical Research Involving Animals. The experiments were approved by the departmental veterinary services of Paris (approval number: A75 14-02).

Animals

The generation of the recombinant substrains from 66HMMU1 at the Pasteur Institute (Paris) has been previously reported (Burgio et al, 2007; Vatin et al, 2012). After weaning, 4 weeks aged mice were maintained in an animal facility of the Cochin Institute (Paris) at normal temperature (21–23°C) and 14 h light/10 h dark photoperiods with free access to water and food. For all experiences, animals were killed by cervical dislocation.

Phenotyping by high frequency ultrasonography

To evaluate the implantation rate, mice from B6 (reference strain) and Rc strains were crossed and phenotyped at the small animal imaging facility of the Cochin Institute using high frequency ultrasonography (VEVO 770, Visulasonics, Toronto, Canada). Each female was used one time to collect phenotypic data from the primo-gestation. Briefly, a chemical hair remover was used to eliminate abdominal hair. Ultrasonographic contact gel was used to ensure contact between the skin surface and the transducer. Body temperature, electrocardiographic and respiratory profiles were monitored using ultrasound device's integrated heating pad and monitoring device (THM150, Indus Instruments, Webster, TX, USA). The implantation rate was determined early in the gestation, at E7.5 and E9.5. At these stages, the small size of embryos permits a fluent count and resorbed embryos are also visible.

Evaluation of in vivo fertilization

B6 females were mated with Rc3 or B6 males on the night (a couple per cage) and the next day females with a positive plug were isolated. Oocytes were collected from their ampullae of oviducts and freed from the cumulus cells by brief incubation at 37°C with hyaluronidase in M2 medium. Oocytes were rinsed and kept in M2 medium in a humidified 5% CO2 atmosphere at 37°C before their mounting with Vectashield/DAPI. The number of fertilized oocytes was evaluated using 4′, 6-diamidino-2-phenylindole (DAPI) immune-fluorescent staining to visualize the DNA.

In vitro fertilization

Oocytes Preparation: B6 females (5 weeks old) were injected with 5 U of PMSG (pregnant mare's serum gonadotrophin) (Intervet, France) followed by 5 U of hCG (human chorionic gonadotrophin) (Intervet) 48 h later. After the super-ovulation, cumulus-oocyte complexes

were collected from ampullae of oviducts 14–16 h after hCG injection. Oocytes were freed from the cumulus cells by 3-5 min incubation at 37°C with hyaluronidase (Sigma) in M2 medium (Sigma). The oocytes were rinsed and kept in M2 medium at 37°C under 5% CO2 atmosphere under mineral oil. When experiments were performed with zona-free oocytes, ZP was then dissolved with acidic Tyrode's (AT) solution (pH 2.5, Sigma) under visual monitoring. The zona-free eggs were rapidly washed in M2 medium and kept at 37°C under 5% CO2 atmosphere for 2 to 3 hr to recover their fertilization ability.

Sperm preparation: Mouse spermatozoa were collected from the caudae epididymis of B6 and Rc3 males (8–13 weeks old) and capacitated at 37°C under 5% CO2 for 90 min in a 500 ml drop of Ferticult medium (FertiPro, Belgium) supplemented with 3% BSA (Sigma), under mineral oil.

In Vitro Fertilization: Cumulus-intact and zona-free eggs were inseminated with capacitated sperm for 3 hr in a 100 µl drop of Ferticult 3% BSA medium at a final concentration of 10^6/ml or 10^5/ml, respectively. Then, they were washed and directly mounted in Vectashield/DAPI for microscopy observation. The oocytes were considered fertilized when they showed at least one fluorescent decondensed sperm head within their cytoplasm.

Assessment of Rc3 sperm head morphology

Freshly recovered or capacitated sperm were washed in PBS containing 1% BSA, centrifuged at 300 g for 5 minutes and immediately fixed in 4% PFA in PBS-1% BSA at 4°C for 1 hr. After washing, the fixed spermatozoa were submitted to the PSA-staining protocol for the detection of acrosomal status. Spermatozoa were treated for 30 minutes in 95% ethanol at 4°C, washed by centrifugation through PBS and stained with FITC-conjugated lectin PSA (25 µg/ml in PBS) for 10 minutes. The nucleus was stained in red with Ethidium bromide. After repeated washing with double distilled water, a drop of sperm suspension was smeared on slide, air-dried, mounted with Vectashield.

Suspension cell sorting

Cells from 2-month old mice were obtained as described previously (Lassalle et al, 2003). The albuginea were removed and the seminiferous tubules were dissociated using enzymatic digestion by collagenase type I at 100u/ml for 15 minutes at 32°C in HBSS supplemented with 20 mM HEPES pH 7.2, 1.2 mM MgSO47H2O, 1.3 mM CaCl22H2O, 6.6 mM sodium pyruvate, 0.05% lactate. After an HBSS wash and centrifugation, the pelleted tubules were further incubated in Cell Dissociation Buffer (Invitrogen) for 25 minutes at 32°C. The resulting whole cell suspension was successively filtered through a 40 µm nylon mesh and through a 20 µm nylon mesh to remove cell clumps. After an HBSS wash, the cell pellet was resuspended in incubation buffer (HBSS supplemented with 20 mM HEPES pH 7.2, 1.2 mM MgSO47H2O, 1.3 mM CaCl22H2O, 6.6 mM sodium pyruvate, 0.05% lactate, glutamine and 1% fetal calf serum) and stained with Hoechst 33342 (5µg/ml) for 1h at 32°C in a water bath. Cells were then labeled with monoclonal antibodies (1 µg per 10^6 cells) from BD

Pharmingen: anti-c-Kit-biotin (2B8), and anti-α-6 integrin-PE (GoH3). Cell-sorting were performed on ARIA (Becton Dickinson).

RNA extraction and Quantitative RT-PCR

Total RNA was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. After RNA preparation, the total RNA was treated with DNase I (Invitrogen Life Technologies) for 10 min at room temperature followed by inactivation with EDTA (Sigma-Aldrich). Total RNA was reverse transcribed to obtain cDNA using M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) following manufacturer's protocols. Quantitative PCR was carried out using fast SYBR Green Master Mix (Applied Biosystems) and a real time PCR system (Light Cycler 1.5, Roche Diagnostics, Division Applied Sciences, Meylan, France) according to standard PCR conditions. To validate the primers used in qRT-PCR, four pairs of primers were tested for each candidate gene and reference gene. For quantitative calculations, values were normalized to mouse *Cyclophiline A* expression.

Statistical Analysis

Results are presented as the mean ± SEM calculated from the variation between individual female. The statistical significance of the differences observed between Rc strains and the B6 control was evaluated by t-test using the Bonferroni-corrected levels. Statistically significant results are labeled as follows in all figures: *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$.

Results and Discussion

In vivo phenotyping:

Previously we have presented the generation by recombination of the IRCS substrains (Rc) from 66HMMU1 (Vatin et al. 2012). Each of them harbored one spretus fragment that overlap with those of the other substrains. These substrains differ from B6 (control) by the presence in there genome of a unique spretus fragment on the chromosome 1 (Figure 1). Here we present the phenotyping of two substrains Rc3 and Rc4 by the use of an in vivo ultrasonic method to evaluate the implantation rate. In this aim, we performed homozygous crosses with males and females from each substrain and determined the mean of implanted embryos. The results were compared to those from the homozygous cross B6 control (Figure 2). We observed a significant diminution of the mean of implanted embryos in the Rc3 crosses compared to B6 control ($p=0.0003$). To know whether the defect observed in Rc3 concern the female or the male side, we performed two heterozygous crosses: ♂Rc3x♀B6 and ♂B6x♀Rc3.We found the phenotype only in the cross ♂Rc3x♀B6 ($p=0.0002$) and we concluded that the defect was in the male side (Figure 2).

QTL fine mapping:

The Rc3 substrain differs from the B6 control by his spretus fragment of 10 Mb on the MMU1 chromosome delimited by D1Mit438 and Rs3692309 markers (see Figure 1). This spretus region containing genes responsible for the observed phenotype, but in the aim to identify the minimal QTL region we performed a genotype/phenotype segregation. We therefore excluded the *spretus* region of 4Mb between the two markers D1Mit440 and Rs3692309 shared by Rc3 and Rc4 that not have the phenotype of interest (see Figure 1). Furthermore, to exclude the hypothesis of putative epistatic interaction, we also analyzed the phenotype of the Rc1 substrain. This last strain shared a large spretus region with Rc3 (see Figure 1). It's also exhibited the phenotype ($p = 0.034$ compared to B6 control) but the genetic difference between Rc3 and Rc1 not leads to difference on their phenotype ($p= 0.17$) (Figure 3). This observation exclude the influence of the region delimited by Rs622012 and Rs3692309 markers (see Figure 1). We can define a QTL region of about 6Mb responsible for a diminution of the litter size on the chromosome 1 between D1Mit438 and D1Mit440 (hatched zone in Figure 1). Using the BioGPS database we identify 4 genes hightly and specifically expressed in the testis : *dnajb3*, *efhd1*, *spata3* and *pde6d*.

Phenotype characterization:

To identify the origin of the male's defect, we performed matings using Rc3 males and B6 females. The next day, we collected and count the number of fertilized oocytes from ampullae of oviducts of females with a positive plug to investigate in vivo fertilization (figure 4). We compared the results from Rc3 with those from B6 control and observed a significant diminution of fertilization rate (FR) with Rc3 male ($p= 0,039$). We also performed in vitro fertilization using B6 oocytes and Rc3 epididymal spermatozoa and we observed a significant diminution of their fertilizing ability compared to B6 control in terms of Fertilization Index (FI) ($p=2,4*10^{-15}$ and $2,5*10^{-14}$ for FR and FI respectively). Interestingly, when we increased the concentration of Rc3 spermatozoa (10 times), the fertilization rate and index were

corrected (Figure 5). Indeed, we estimated that in Rc3 sperm there are approximately 80% of immobile sperm versus 15 to 20% in B6 sperm.

All this results suggest a dysfunction of Rc3 sperm. To check the origin of this dysfunction, and after an inconclusive observation under phase-contrast microscopy, Rc3 spermatozoa were observed using fluorescence microscopy with PSA-FITC to analyze sperm acrosome. In a significant percentage of sperm, the acrosome was reduced and sometimes completely absent even though sperm were labeled immediately after their recovery, before they do their capacitation (Figure 6). A second observation of Rc3 spermatozoa using another method, electron microscopy, confirmed the acrosome defect (Figure 7) but also reveled anomalies of the flagella (figure 8). Indeed in Rc3 sperm, there is a residual body at flagella not found on B6 sperm. In addition, these residual bodies contain large lipid droplets that likely reflect a metabolic dysfunction. This morphological abnormalities seems to characterize a severe teratozoospermia, namely globozoospermia.

Identification of candidate genes in the QTL region responsible for globozoospermia

Globozoospermia is a rare and severe teratozoospermia characterized by round-headed spermatozoa lacking an acrosome, an important actor playing a crucial role during fertilization (Dam AH, Koscinski I,, 2007). Also maturation defects such as persisting residual cytoplasmic body/droplet surrounding the nucleus or midpiece have frequently been reported (Dam AH, Koscinski I,, 2007). Similar pathogenesis is found in both mice and men. These morphological anomalies leading to globozoospermia, suggest that the defect occurs during the last step of spermatogenesis. The pathogenesis of globozoospermia most probably originates during acrosome formation and head elongation. We can propose that the causal gene is crucial for the spermatid-spermatozoa transition and maturation. The expression level of the four genes (dnajb3, efhd1, spata3 and pde6d) in the different cellular population of the testis was investigated using RT-PCR. Three of them are expressed in the cellular population of interest, spermatids: dnajb3, efhd1 and spata3 (Figure 9). To date, mutations in two genes *SPATA16* (also named NYD-SP12) and *DPY19L2* have been identified as responsible for globozoospermia. Interestingly, in our study, we have identified a QTL region on mouse chromosome 1 containing gene among them *Spata3* is a member of the spermatogenesis-associated family. Like *SPATA16* (Xu M, Xiao J, 2003), we demonstrated that *Spata3* is specifically expressed in mouse testis in spermatocyte and spermatids. These are the first elements of this gene as a good candidate to explain the phenotype.

Figures legends:

Figure 1: Genotypes of the different recombinant strains (Rc) and the parental strains (66HMMU1 (SEG) and B6).

Recombinant strains (Rc) were generated at the Pasteur institute (Paris) from 66HMMU1 strain by recombination events inside the MMU1 spretus segment. Blue regions correspond to B6 background and yellow regions to spretus fragments on chromosome 1. Marker positions are given in mega base pairs (Mbp).

Figure 2: Ultrasonographic examination of the implantation in different crosses using mice from Rc3, Rc4 and B6 strains.

Different crosses were perfomed using males and females from Rc3, Rc4 and B6 strains (three homozygous crosses using Rc3, Rc4 and B6 strains and two inverse heterozygous crosses with Rc3 and B6 strains). The results are presented as the mean (± SEM) of implanted embryos for (n) gestations.

Figure 3: Ultrasonographic phenotyping of the Rc3 and Rc1 strain on the implantation rate.

Homozygous crosses were performed using the Rc3, Rc1 and B6 strain and the number of implanted embryos were determined. The results are prensented as the mean of implanted embryos (± SEM) for (n) gestations.

Figure 4: Evaluation of in vivo fertilization rate.

B6 females were mated with B6 or R3 males to evaluate their respective capacity to fertilize oocytes. The mating was performed on the night and the next day, oocytes from ampullae of oviducts of females with a positive plug were collected. The fertilization was assessed using DAPI immune-fluorescent staining. The results are presents as the mean of fertilization rate ± SEM.

Figure 5: In vitro fertilization (IVF) using B6 or Rc3 sperm

Fertilization rate (or mean ± SEM of fertilized eggs) and fertilization index (or mean ± SEM of sperm number fused by egg) on zona-free IVF assay at 10^5 spermatozoa per ml for 3 hr. The experiments were repeated three times.

Figure 6: Morphology of B6 (A) and Rc3 (B) epididymal sperm observed by fluorescence microscopy.

The nucleus (N) was stained in red with Ethidium bromide and the acrosomal region (a) in green with PSA-FITC. We noted reduced acrosome and sometimes its total absence in some Rc3 sperm.

Figure 7: Observation of B6 (A and B) and Rc3 (C to E) sperm using electron microscopy.

Electron microscopy examination revealed acrosomal defect in Rc3 spermatozoa absent in B6 sperm.

Figure 8: Observation of B6 (A) and Rc3 (B to D) flagella using electron microscopy

Electron microscopy examination revealed anomalies in the flagella of Rc3 spermatozoa, as lipoid drop. This phenomenon is present only on Rc3 spermatozoa and never observed in B6.

Figure 9: Expression level of the four candidate genes during the different steps of spermatogenesis.

To determine the step of spermatogenesis in which the four candidate genes are expressed, the different cellular population were separated and the expression level were measured by RT-PCR.

Figures:

Figure 1:

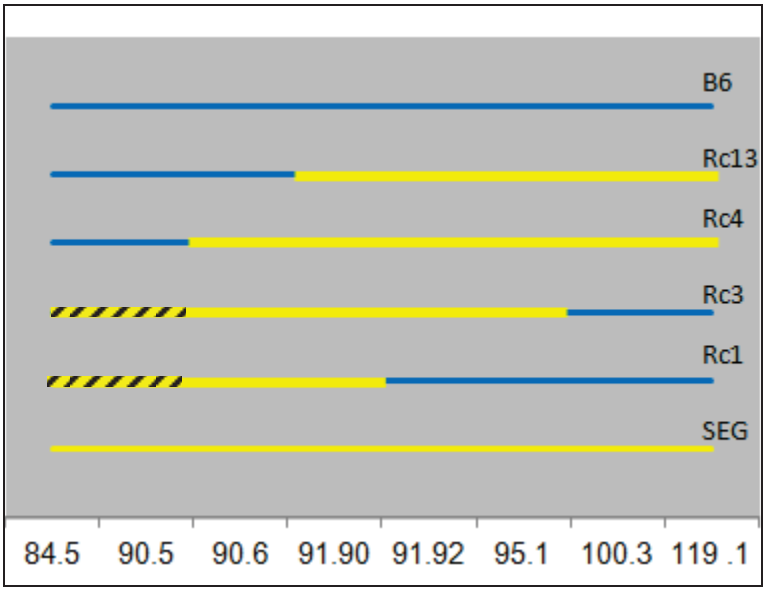


Figure 2:

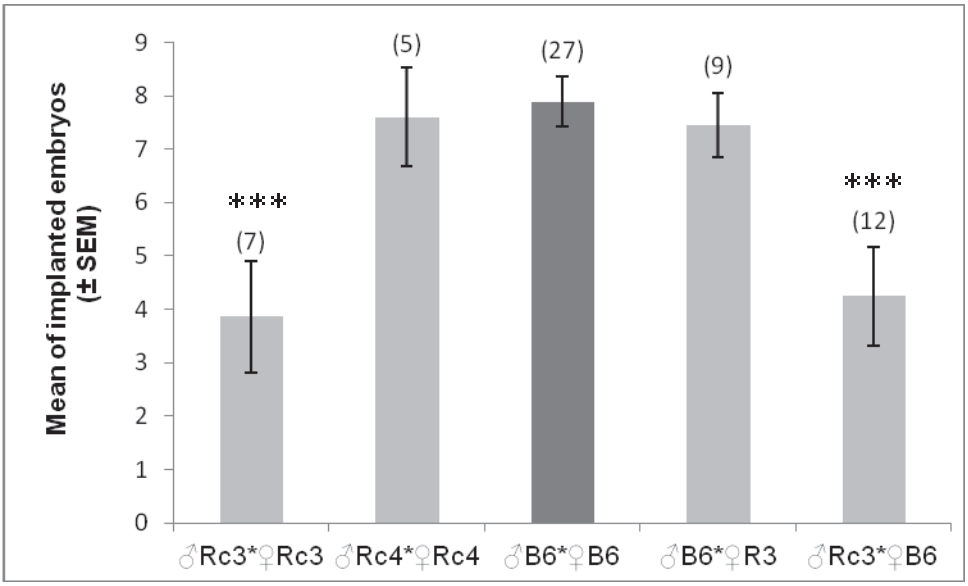


Figure 3:

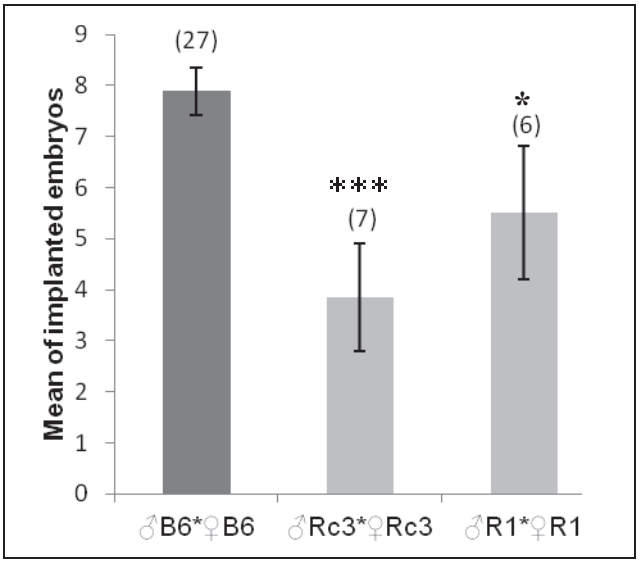


Figure 4:

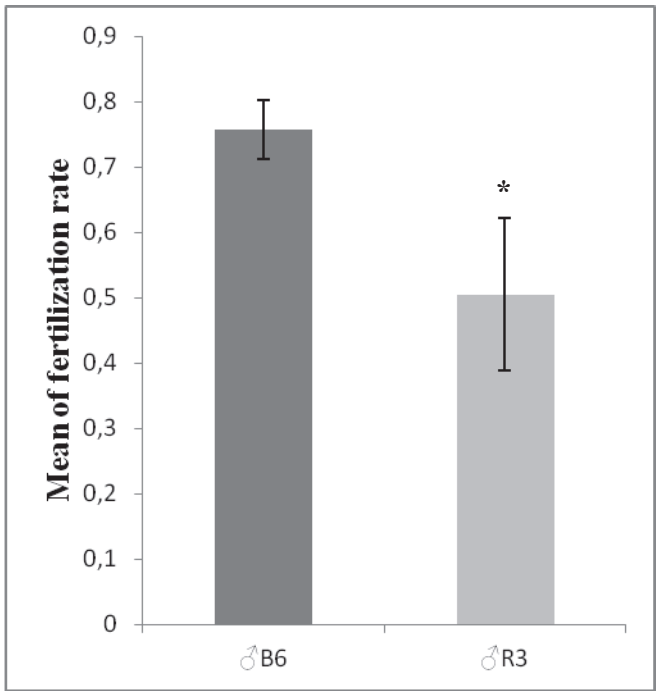


Figure 5

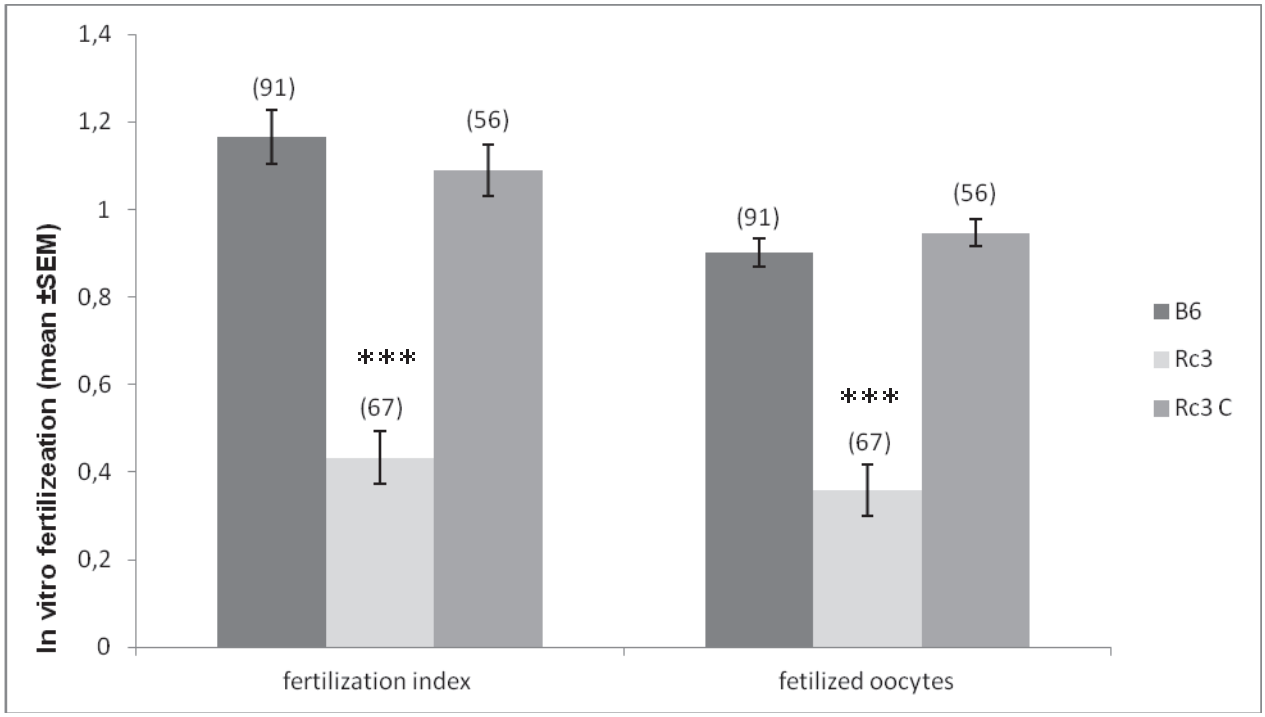


Figure 6:

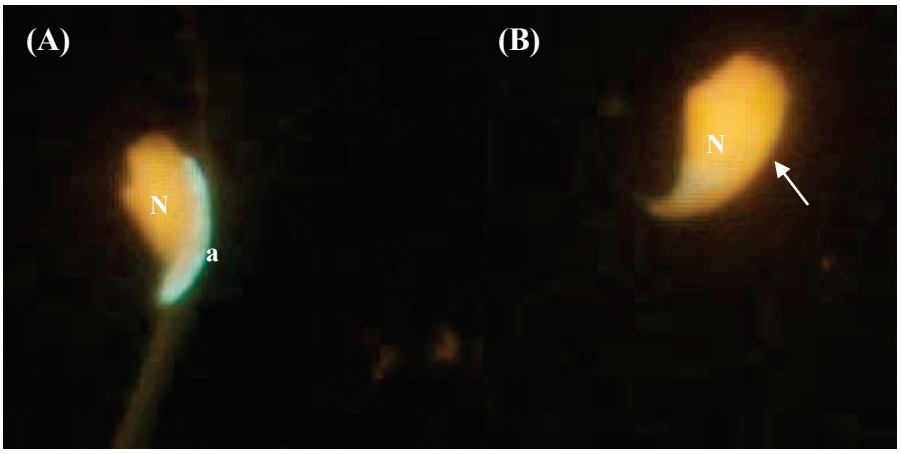


Figure 7

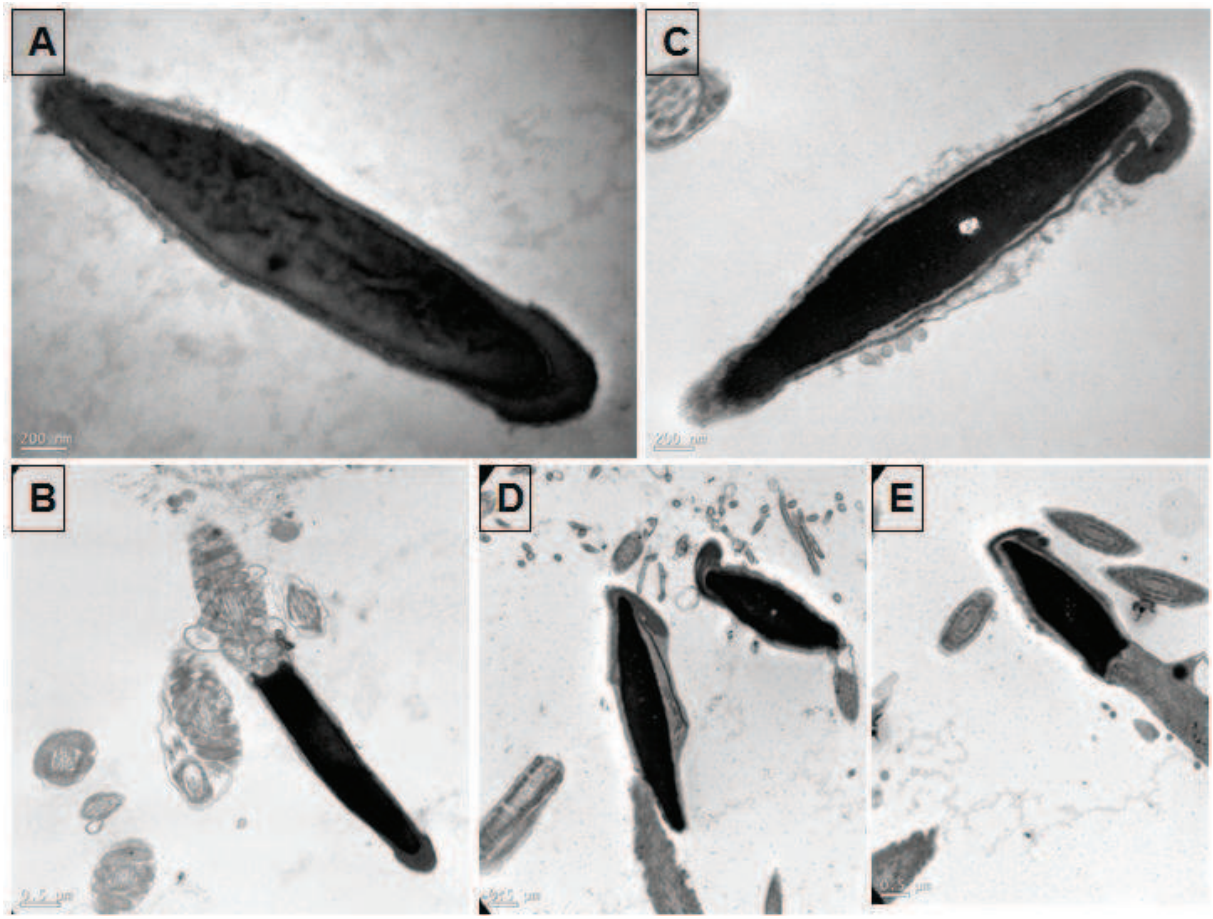


Figure 8

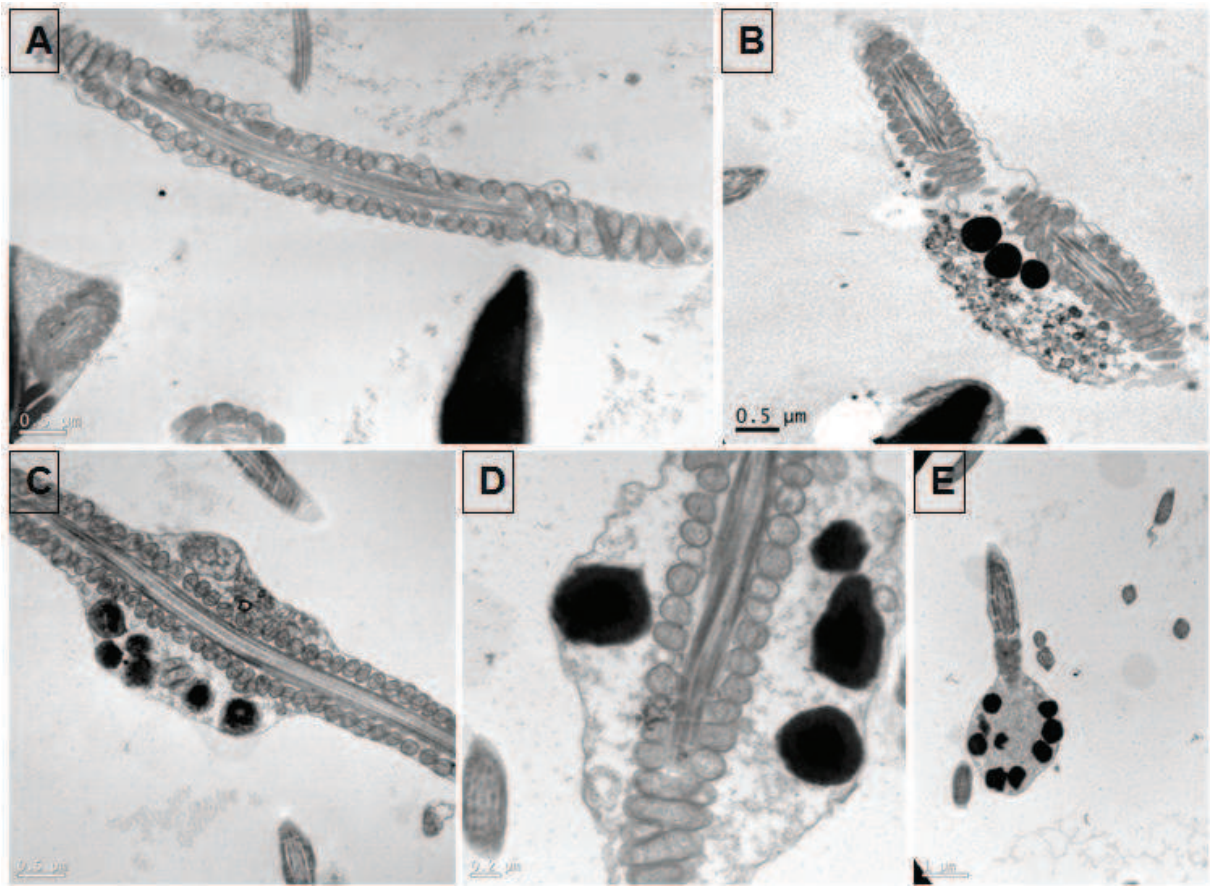
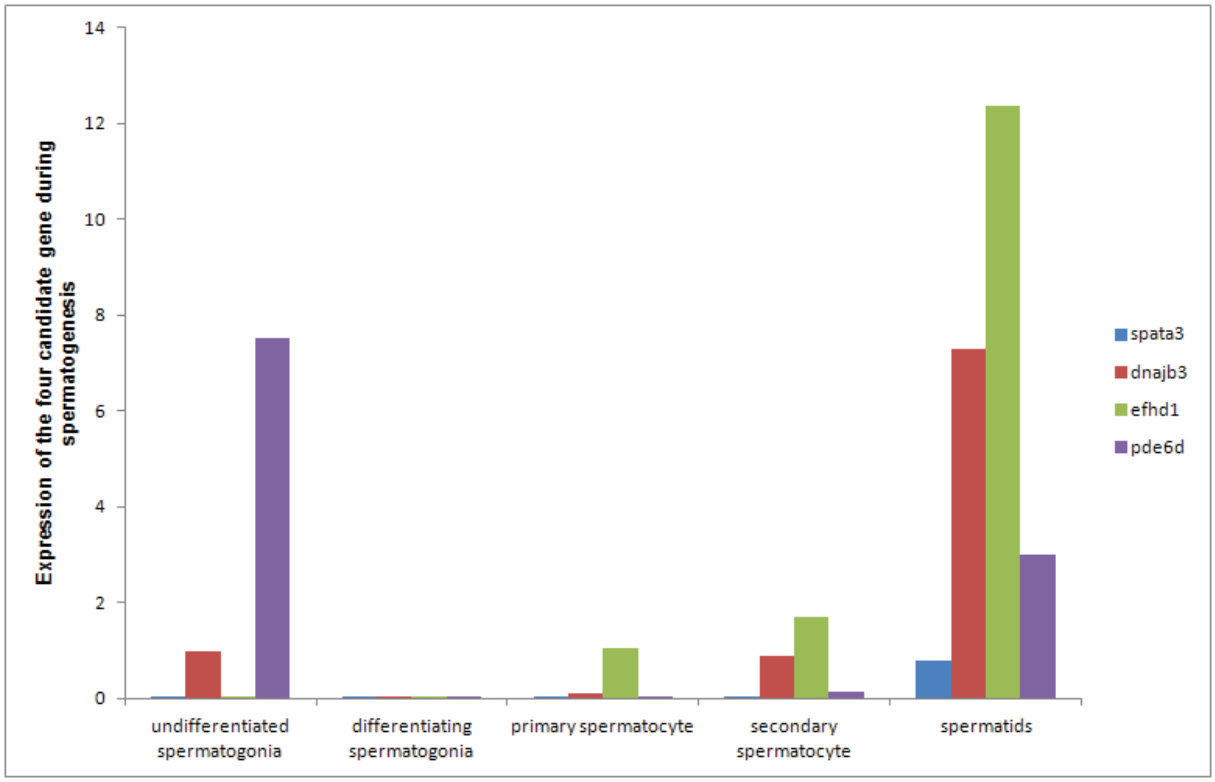


Figure 9


Expression of the four candidate gene during spermatogenesis
14
12
10
8
6
4
2
0
undifferentiated spermatogonia
differentiating spermatogonia
primary spermatocyte
secondary spermatocyte
spermatids
spata3
dnajb3
efhd1
pde6d

References:

Richard M. Sharpe, Chris McKinnell, Catrina Kivlin and Jane S. Fisher. Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood Reproduction. 2003 Jun;125(6):769-84.

O'Donnell L, Nicholls PK, O'Bryan MK, McLachlan RI, Stanton PG. Spermiation: The process of sperm release. Spermatogenesis. 2011 Jan;1(1):14-35.

Cheng CY, Mruk DD. The biology of spermatogenesis: the past, present and future. Philos Trans R Soc Lond B Biol Sci. 2010 May 27;365(1546):1459-63.

Borg CL, Wolski KM, Gibbs GM, O'Bryan MK. Phenotyping male infertility in the mouse: how to get the most out of a 'non-performer'.Hum Reprod Update. 2010 Mar-Apr;16(2):205-24.

Schultz N, Hamra FK, Garbers DL. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets.Proc Natl Acad Sci U S A. 2003 Oct 14;100(21):12201-6.

Massart A, Lissens W, Tournaye H, Stouffs K. Genetic causes of spermatogenic failure.Asian J Androl. 2012 Jan;14(1):40-8.

Venables JP, Cooke HJ. Lessons from knockout and transgenic mice for infertility in men.J Endocrinol Invest. 2000 Oct;23(9):584-91.

Escalier D. Impact of genetic engineering on the understanding of spermatogenesis. Hum Reprod Update. 2001 Mar-Apr;7(2):191-210.

Cooke HJ, Saunders PT. Mouse models of male infertility. Nat Rev Genet. 2002 Oct;3(10):790-801.

Fox MS, Ares VX, Turek PJ, Haqq C, Reijo Pera RA. Feasibility of global gene expression analysis in testicular biopsies from infertile men. Mol Reprod Dev. 2003 Dec;66(4):403-21.

Burgio G, Szatanik M, Guénet JL, Arnau MR, Panthier JJ, Montagutelli X. Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. Genetics. 2007 Dec;177(4):2321-33.

L'Hôte D, Serres C, Laissue P, Oulmouden A, Rogel-Gaillard C, Montagutelli X, Vaiman D. Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. Genetics. 2007 Jul;176(3):1907-21.

L'Hôte D, Laissue P, Serres C, Montagutelli X, Veitia RA, Vaiman D. Interspecific resources: a major tool for quantitative trait locus cloning and speciation research. Bioessays. 2010 Feb;32(2):132-42.

Vatin M, Burgio G, Renault G, Laissue P, Firlej V, Mondon F, Montagutelli X, Vaiman D, Serres C, Ziyyat A. Refined mapping of a quantitative trait locus on chromosome 1 responsible for mouse embryonic death. PLoS One. 2012;7(8):e43356.

Dam AH, Koscinski I, Kremer JA, Moutou C, Jaeger AS, Oudakker AR, Tournaye H,

Charlet N, Lagier-Tourenne C, van Bokhoven H, Viville S. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. Am J Hum Genet. 2007 Oct;81(4):813-20.

Xu M, Xiao J, Chen J, Li J, Yin L, Zhu H, Zhou Z, Sha J. Identification and characterization of a novel human testis-specific Golgi protein, NYD-SP12. Mol Hum Reprod. 2003 Jan;9(1):9-17.

Discussion

Le privilège de se marier et de fonder une famille est considéré comme un droit humain fondamental (Organisation des Nations Unies, 1948). Cependant, 15% des couples présentent des problèmes d'hypofertilité voire de stérilité. Des facteurs génétiques mais également des facteurs liés au mode de vie sont en cause dans l'augmentation de l'infertilité observée dans les pays industrialisés. Ces derniers incluent l'âge des femmes, le tabagisme, le poids, l'alimentation, l'exercice, le stress psychologique, la consommation de caféine, la consommation d'alcool et l'exposition à des polluants environnementaux (Homan et al., 2007). Par exemple, l'augmentation de la prévalence de l'obésité est associée à une anovulation et au syndrome des ovaires polykystiques (Nelson and Fleming, 2007), l'incidence des maladies sexuellement transmissibles affectent les organes reproducteurs, par exemple, les cas d'infection par chlamydia (Low and Ward, 2007) et le recul de l'âge de la conception du premier enfant peut entraîner une infertilité (Commission des Communautés Européennes, 2005, 2006). L'augmentation des cas d'infertilité se traduit par une hausse croissante du recours aux techniques de reproduction assistée (Nybo Andersen et al., 2004; Nygren and Andersen, 2001a; Nygren and Andersen, 2001b). La reproduction assistée regroupe «tous les traitements ou les procédures qui comprennent la manutention in vitro des ovocytes, des spermatozoïdes et des embryons humains dans le but d'établir une grossesse

(Zegers-Hochschild et al., 2006a). La grande souffrance psychologique des couples qui consultent, le coût économique très élevé des traitements et des interventions (pris en charge à 100 % par la sécurité sociale), les nombreuses questions éthiques soulevées par les pratiques médicales dans ce domaine font de la reproduction un enjeu de santé publique majeur dont l'étude des défauts semble incontournable.

L'étude de la génétique des maladies complexes est délicate car le phénotype observé est le résultat d'interactions entre de nombreuses protéines, codées par des gènes différents. Par conséquent, il est indispensable de diversifier les approches afin de découvrir les différents acteurs. Ainsi, au laboratoire, nous avons choisi un modèle d'étude original, les lignées murines interspécifiques recombinantes congéniques (IRCS) créées à l'Institut Pasteur par de Dr Montagutelli. Dans ce modèle, environ 2% du génome *spretus* est dispersé dans le fond génétique B6 (Burgio et al., 2007; L'Hote et al., 2007).

Dans le cadre d'une collaboration avec le Dr Montagutelli, l'exploitation du modèle IRCS pour la recherche de gènes candidats responsables de défaut de fertilité humaine a été initiée en 2005 au sein de l'équipe du Dr Daniel VAIMAN. Ces travaux ont permis la réalisation de 3 thèses dont celle qui fait l'objet du présent manuscrit et ont permis la publication de plusieurs articles

scientifiques avec d'autres travaux en cours de finalisation pour publication. Il s'agissait en effet d'un travail laborieux de phénotypage à grande échelle (organes reproducteurs, gamètes, fécondation, implantation et développement embryonnaire), suivi de la cartographie fine de QTL puis du clonage positionnel de gènes. Ces travaux ont toutefois été fructueux car, ils ont permis la localisation de 4 QTL dans lesquels des gènes candidats ont pu être identifiés (L'Hôte et al., 2011; Laissue et al., 2009; Vatin et al., 2012).

Au cours de ma thèse je me suis intéressée aux causes génétiques d'infertilité masculine liée à des défauts de spermatogenèse. La régulation du processus complexe de la spermatogenèse dépend de la coopération de nombreux gènes qui sont exprimés à différents stades de différenciation, de division mitotiques et méiotiques, et de développement post-méiotique conduisant finalement à des spermatozoïdes (Sharpe et al., 1994).

J'ai donc pu participer au clonage positionnel d'un gène capable de moduler le poids des testicules en interférant avec la méiose. Ce gène *Fignl1* est localisé sur le chromosome 11 chez la souris et code pour une protéine appartenant au groupe des AAA-ATPases associées aux diverses activités cellulaires. Par les travaux réalisés au laboratoire, *Fignl1* a été montré comme étant principalement

exprimé au niveau des cellules germinales méiotiques, les spermatocytes. Son rôle, en particulier dans le passage de la mitose à la méiose chez *C. elegans* (Luke-Glaser et al., 2007) et l'existence de SNP non synonymes entre *spretus* et *musculus* laisse penser que Fignl1 pourrait jouer un rôle déterminant dans la régulation de la durée de la méiose. Dans notre modèle, Fignl1 en version *spretus* a probablement des difficultés à être pris en charge par le système d'ubiquitinylation *musculus* du fait du polymorphisme entre SEG et B6. Une autre différence entre le Fignl1 *spretus* et *musculus* est l'existence d'une version *spretus* tronquée supplémentaire correspondant à un variant d'épissage pouvant rentrer en compétition avec le variant de grande taille dans le processus de dégradation. C'est donc l'accumulation de la protéine capable de ralentir le processus de la méiose se traduisant par une apoptose et une fréquence accrue d'anomalies des spermatozoïdes, qui est à l'origine du phénotype observé. Cependant la génération de souris KO pour le gène ou la substitution de *Fignl1 spretus* dans un contexte génomique B6, constituerait la preuve ultime de l'implication du gène. En effet, les études utilisant des modèles animaux de souris transgéniques invalidées et présentant un phénotype de stérilité, continueront à nous apporter, dans les années à venir, des informations précieuses sur d'autres gènes impliqués dans l'infertilité.

D'autre part, dans le cadre de travaux récents, j'ai réalisé le phénotypage des lignées recombinantes obtenues à partir de 66HMMU1 par échographie à haute fréquence et j'ai observé une diminution significative de la moyenne du nombre d'embryons implantés dans le croisement homozygote de la lignée IRCS Rc3 comparée au croisement contrôle homozygote B6. J'ai démontré que le phénotype observé résulte d'un défaut de fertilité des mâles Rc3, à savoir une globozoospermie partielle. En effet, l'observation de la morphologie des spermatozoïdes a révélé une anomalie au niveau de la tête : une réduction de la taille de l'acrosome voire son absence sur une partie de la population spermatique. Les premières observations en microscopie électronique semblent confirmer ces résultats relatifs au défaut d'acrosome mais révèlent également la présence en grand nombre de gouttelettes lipidiques au niveau du flagelle. La région QTL en cause a été cartographiée finement sur le chromosome 1. Dans le but de proposer des gènes candidats susceptibles d'être à l'origine de cette hypofertilité, la même approche que celle des précédentes études est employée en se basant sur leur niveau d'expression testiculaire, sur la présence de SNP entre SEG et B6 et sur les données de la littérature. Ces travaux sont actuellement en cours de réalisation.

La recherche sur l'infertilité masculine et la variété de causes possibles ont déjà permis de mieux comprendre le processus de la spermatogenèse et de la fertilité masculine. Une meilleure connaissance des mécanismes de la spermatogenèse, sans négliger le rôle de l'environnement, devrait permettre de trouver de nouvelles solutions thérapeutiques. L''identification des gènes potentiellement impliqués dans la spermatogenèse doit être poursuivie afin d'accroître l'efficacité des tests de diagnostic indispensables pour envisager un traitement médical approprié des individus stériles. Le « microarray » est un outil puissant de recherche permettant d'extraire des informations importantes du génome des mammifères. Par exemple, en 2012, Waclawska et Kurpisz ont présenté l'utilisation de la technique de puces pour identifier des gènes critiques de la spermatogenèse afin de réaliser une corrélation avec des cas d'infertilité (Waclawska and Kurpisz, 2012). Aux niveaux pré-méiotique et méiotique, des gènes de la famille *DAZ* et *TSPY1* sont fortement exprimés, alors qu'au cours de la phase post-méiotique, l'expression des gènes *PRM2, TNP1, SYNJ2, et ZPBP* augmente. Ces gènes sont principalement responsables de la formation de la tête des spermatozoïdes et du fonctionnement normal de ces derniers (Fox et al., 2003). Chez les patients atteints de la microdélétion AZFc, quelques gènes post-méiotiques comme *TNP1* et *PRM2* sont faiblement exprimés dans le groupe de patients caractérisés par un faible nombre de cellules post-méiotiques, d'où

l'indication de leur importance à ce stade de la spermatogenèse (Gatta et al., 2010).

L'objectif ultime est d'identifier des biomarqueurs potentiels utilisables dans le cadre de plateforme de diagnostic afin de permettre l'application de la thérapie médicale appropriée. Ainsi, il semble possible de simplifier les tests de diagnostic pour la stérilité en utilisant les profils de l'expression de gènes (Moldenhauer et al., 2003). Le gène *Fignl1* pourrait être étudié dans le cadre de stérilité humaine avec arrêt idiopathique de la méiose.

Au cours de mon travail de thèse, je me suis également intéressée à une pathologie de la grossesse. Un à 5% des couples sont concernés par les fausses couches répétées (RSA), définies par l'interruption subite de trois ou plusieurs grossesses consécutives (Rai and Regan, 2006). Les anomalies chromosomiques dans l'embryon sont la cause la plus fréquente tandis que les facteurs utérins sont invoqués pour expliquer les RSA non liées à une anomalie chromosomique. Ces facteurs utérins sont cependant mal définis (Teklenburg et al., 2010). Au laboratoire, deux régions QTL ont été identifiées comme étant associées à une augmentation des cas de mort embryonnaire chez la souris (Laissue et al., 2009). La première au niveau du chromosome 1 correspond à un intervalle d'une

trentaine de Mb contenant plus de 200 gènes. En utilisant une approche de phénotypage in vivo par échographie à haute fréquence, j'ai pu réaliser la cartographie fine du QTL et proposer une région de 6 Mb contenant une trentaine de gènes. L'analyse génétique suggère que le phénotype observé est lié à un dysfonctionnement utérin. Alors, appliquant plusieurs filtres comme le niveau d'expression dans l'utérus et l'existence de SNP, et m'appuyant sur une recherche bioinformatique j'ai pu proposer 7 gènes candidats (*Ncl*, *Cab39*, *Eif4e2, Trip12, Psmd1, Cops7b and Usp40)* qui ont été retenus comme potentiellement responsables du phénotype observé (Vatin et al., 2012). De plus, les candidats retenus codent pour des protéines appartenant à différentes voies de signalisation telles que la voie mTOR ou celle de la dégradation des protéines dépendante de l'ubiquitine/protéasome, et sont impliqués dans plusieurs processus physiologiques tels que l'angiogenèse.

Le second QTL correspond à une région de 3Mb environ contenant une trentaine de gènes, localisée sur le chromosome 13 murin. Son étude, combinant une approche de clonage positionnel, l'analyse transcriptomique et des tests in vitro, a permis de valider pour la première fois le clonage positionnel d'un QTL par génotypage dans l'espèce humaine. Nos résultats montrent que le facteur de transcription Foxd1, qui module l'expression du facteur de croissance

placentaire Pgf et du facteur C3 du complément, est un nouvel acteur moléculaire majeur nécessaire au bon déroulement de la grossesse. En effet, en collaboration avec le Pr Gris du CHU de Nîmes, par le séquençage de *FOXD1* chez des patientes atteintes de RSA idiopathiques et chez des femmes contrôles, nous avons identifié des variants *FOXD1* exclusivement chez les patientes atteintes de RSA (Laissue et al, en préparation).

Les études sur le modèle murin nous ont également amené à nous intéresser au gène *Alpp* codant pour l'alkaline phosphatase placentaire. J'ai mis au point et réalisé le séquençage de 100 patientes RSA idiopathiques et 100 contrôles d'origine méditerranéenne. Cette étude préliminaire a permis de mettre en évidence des variations de séquence dans le gène susceptibles d'expliquer environ 10% des cas d'avortements spontanés répétés dans la population étudiée. Des SNP ont été identifiés comme significativement associés à des cas de RSA parmi lesquels le polymorphisme rs130226692 (Ile89Leu) Une analyse en 3D de la structure de la protéine montre que le résidu Ile89 est sur la surface de la molécule et à proximité de l'interface homodimèrique. Il est à noter que les résidus dans la même zone ont été montrés comme étant indispensables pour éviter la dimérisation entre isoformes paralogues. De plus, un effet de la substitution Ile89Leu sur la stabilité du dimère, par exemple, ne peut pas être

exclu. Des études antérieures menées en Europe du Nord (Finlande et Suède) ont déjà associé des polymorphismes d'ALPP avec les RSA, cependant nous n'avons pas trouvé, dans notre population méditerranéenne, les SNP identifiés dans l'étude d'Europe du Nord, sans doute en raison de l'origine géographique des populations (Beckman et al., 1972).

La plupart des études d'association génétique ciblant les RSA ont été conçues comme des hypothèses fondées sur des études de gènes candidats. Les gènes les plus souvent étudiés sont associés à l'immunotolérance. Les polymorphismes dans près de 90 gènes différents ont été étudiés, cependant dans la plupart des études, l'association entre un polymorphisme et la pathologie est négative. En clinique, les polymorphismes les plus souvent testés chez des patientes RSA sont les mutations de la prothrombine (G20210A) et celle de l'enzyme méthylène tétrahydrofolate réductase (C667T) (Rull et al., 2012).

Aussi, nous proposons que la mutation FOXD1 (p.429AlaAla) peut être considérée comme un marqueur moléculaire des RSA qui peut facilement être testé par PCR / séquençage et pourrait être utilisé comme un marqueur moléculaire de la maladie (Laissue et al, en préparation).

En conclusion, nos études confirment le fait que le modèle IRCS est très utile pour la compréhension des maladies multifactorielles complexes chez l'Homme. En effet, l'exploitation des lignée IRCS nous a permis de positionner de nouveaux QTL de fertilité, mais surtout de cloner de nouveaux gènes impliqués dans la fertilité. Ces études pourraient contribuer à élucider les bases moléculaires de certaines infertilités et permettre de proposer de nouveaux marqueurs diagnostiques des différentes pathologies, anomalies de la spermatogenèse et RSA grâce aux progrès dans le séquençage du génome entier et la mise en place d'une base de données des variants «à risque» pour les maladies complexes.

Références bibliographiques

(2008). Genetic aspects of female reproduction. *Hum Reprod Update* **14**, 293-307.

Almeida, A., Layton, M. and Karadimitris, A. (2009). Inherited glycosylphosphatidyl inositol deficiency: a treatable CDG. *Biochim Biophys Acta* **1792**, 874-80.

Andersen, A. N. and Loft, A. (2006). [Use of single embryo in assisted reproduction]. *Tidsskr Nor Laegeforen* **126**, 3082.

Anderson, R. A. and Baird, D. T. (2002). Male contraception. *Endocr Rev* **23**, 735-62.

Balasch, J., Fabregues, F. and Arroyo, V. (1998). Peripheral arterial vasodilation hypothesis: a new insight into the pathogenesis of ovarian hyperstimulation syndrome. *Hum Reprod* **13**, 2718-30.

Beckman, L., Beckman, G. and Magnusson, S. S. (1972). Relationship between placental alkaline phosphatase phenotypes and the frequency of spontaneous abortion in previous pregnancies. *Hum Hered* **22**, 15-7.

Bellazi, L., Germond, S., Dupont, C., Brun-Heath, I., Taillandier, A., De Mazancourt, P., Dieudonne, M. N. and Mornet, E. (2010). A sequence variation in the promoter of the placental alkaline phosphatase gene (ALPP) is associated with allele-specific expression in human term placenta. *Placenta* **31**, 764-9.

Bhasin, S., Ma, K. and de Kretser, D. M. (1997). Y-chromosome microdeletions and male infertility. *Ann Med* **29**, 261-3.

Burgio, G., Szatanik, M., Guenet, J. L., Arnau, M. R., Panthier, J. J. and Montagutelli, X. (2007). Interspecific recombinant congenic strains between C57BL/6 and mice of the Mus spretus species: a powerful tool to dissect genetic control of complex traits. *Genetics* **177**, 2321-33.

Carlborg, O. and Haley, C. S. (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet* **5**, 618-25.

Carreau, S. and Hess, R. A. (2010). Oestrogens and spermatogenesis. *Philos Trans R Soc Lond B Biol Sci* **365**, 1517-35.
Caserta, D., Mantovani, A., Marci, R., Fazi, A., Ciardo, F., La Rocca, C., Maranghi, F. and Moscarini, M. (2011). Environment and women's reproductive health. *Hum Reprod Update* **17**, 418-33.
Cheng, Y. H., Wong, E. W. and Cheng, C. Y. (2011). Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. *Spermatogenesis* **1**, 209-220.
Clermont, Y. and Rambourg, A. (1978). Evolution of the endoplasmic reticulum during rat spermiogenesis. *Am J Anat* **151**, 191-211.
Craig, L. B., Ke, R. W. and Kutteh, W. H. (2002). Increased prevalence of insulin resistance in women with a history of recurrent pregnancy loss. *Fertil Steril* **78**, 487-90.
de La Rochebrochard, E., Soullier, N., Peikrishvili, R., Guibert, J. and Bouyer, J. (2008). High in vitro fertilization discontinuation rate in France. *Int J Gynaecol Obstet* **103**, 74-5.
Dehghani, H., Narisawa, S., Millan, J. L. and Hahnel, A. C. (2000). Effects of disruption of the embryonic alkaline phosphatase gene on preimplantation development of the mouse. *Dev Dyn* **217**, 440-8.
Demant, P. and Hart, A. A. (1986). Recombinant congenic strains--a new tool for analyzing genetic traits determined by more than one gene. *Immunogenetics* **24**, 416-22.
Edmonds, D. K., Lindsay, K. S., Miller, J. F., Williamson, E. and Wood, P. J. (1982). Early embryonic mortality in women. *Fertil Steril* **38**, 447-53.
Empson, M., Lassere, M., Craig, J. C. and Scott, J. R. (2002). Recurrent pregnancy loss with antiphospholipid antibody: a systematic review of therapeutic trials. *Obstet Gynecol* **99**, 135-44.
Evers, J. L. (2002). Female subfertility. *Lancet* **360**, 151-9.
Fauvert, D., Brun-Heath, I., Lia-Baldini, A. S., Bellazi, L., Taillandier, A., Serre, J. L., de Mazancourt, P. and Mornet, E. (2009). Mild forms of hypophosphatasia mostly result from dominant negative effect of severe alleles or from compound heterozygosity for severe and moderate alleles. *BMC Med Genet* **10**, 51.
Ferianec, V. and Linhartova, L. (2011). Extreme elevation of placental alkaline phosphatase as a marker of preterm delivery, placental insufficiency and low birth weight. *Neuro Endocrinol Lett* **32**, 154-7.
Foster, W. G., Neal, M. S., Han, M. S. and Dominguez, M. M. (2008). Environmental contaminants and human infertility: hypothesis or cause for concern? *J Toxicol Environ Health B Crit Rev* **11**, 162-76.

Fox, M. S., Ares, V. X., Turek, P. J., Haqq, C. and Reijo Pera, R. A. (2003). Feasibility of global gene expression analysis in testicular biopsies from infertile men. *Mol Reprod Dev* **66**, 403-21.
Gatta, V., Raicu, F., Ferlin, A., Antonucci, I., Scioletti, A. P., Garolla, A., Palka, G., Foresta, C. and Stuppia, L. (2010). Testis transcriptome analysis in male infertility: new insight on the pathogenesis of oligo-azoospermia in cases with and without AZFc microdeletion. *BMC Genomics* **11**, 401.
Gnoth, C., Godehardt, E., Frank-Herrmann, P., Friol, K., Tigges, J. and Freundl, G. (2005). Definition and prevalence of subfertility and infertility. *Hum Reprod* **20**, 1144-7.
Griswold, M. D., Heckert, L. and Linder, C. (1995). The molecular biology of the FSH receptor. *J Steroid Biochem Mol Biol* **53**, 215-8.
Guenet, J. L. and Bonhomme, F. (2003). Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends Genet* **19**, 24-31.
Hilbert, P., Lindpaintner, K., Beckmann, J. S., Serikawa, T., Soubrier, F., Dubay, C., Cartwright, P., De Gouyon, B., Julier, C., Takahasi, S. et al. (1991). Chromosomal mapping of two genetic loci associated with blood-pressure regulation in hereditary hypertensive rats. *Nature* **353**, 521-9.
Hirahara, F., Andoh, N., Sawai, K., Hirabuki, T., Uemura, T. and Minaguchi, H. (1998). Hyperprolactinemic recurrent miscarriage and results of randomized bromocriptine treatment trials. *Fertil Steril* **70**, 246-52.
Hogge, W. A., Byrnes, A. L., Lanasa, M. C. and Surti, U. (2003). The clinical use of karyotyping spontaneous abortions. *Am J Obstet Gynecol* **189**, 397-400; discussion 400-2.
Homan, G. F., Davies, M. and Norman, R. (2007). The impact of lifestyle factors on reproductive performance in the general population and those undergoing infertility treatment: a review. *Hum Reprod Update* **13**, 209-23.
Jacobs, P. A. and Strong, J. A. (1959). A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature* **183**, 302-3.
Jauniaux, E., Farquharson, R. G., Christiansen, O. B. and Exalto, N. (2006). Evidence-based guidelines for the investigation and medical treatment of recurrent miscarriage. *Hum Reprod* **21**, 2216-22.
Kelly-Weeder, S. and Cox, C. L. (2006). The impact of lifestyle risk factors on female infertility. *Women Health* **44**, 1-23.
Knudsen, U. B., Hansen, V., Juul, S. and Secher, N. J. (1991). Prognosis of a new pregnancy following previous spontaneous abortions. *Eur J Obstet Gynecol Reprod Biol* **39**, 31-6.
Krawitz, P. M., Schweiger, M. R., Rodelsperger, C., Marcelis, C., Kolsch, U., Meisel, C., Stephani, F., Kinoshita, T., Murakami, Y., Bauer, S. et al.

(2010). Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nat Genet* **42**, 827-9.

Kumar, S. (2004). Occupational exposure associated with reproductive dysfunction. *J Occup Health* **46**, 1-19.

Kupferminc, M. J. (2003). Thrombophilia and pregnancy. *Reprod Biol Endocrinol* **1**, 111.

L'Hote, D., Laissue, P., Serres, C., Montagutelli, X., Veitia, R. A. and Vaiman, D. (2010). Interspecific resources: a major tool for quantitative trait locus cloning and speciation research. *Bioessays* **32**, 132-42.

L'Hote, D., Serres, C., Laissue, P., Oulmouden, A., Rogel-Gaillard, C., Montagutelli, X. and Vaiman, D. (2007). Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. *Genetics* **176**, 1907-21.

L'Hôte, D., Vatin, M., Auer, J., Castille, J., Passet, B., Montagutelli, X., Serres, C. and Vaiman, D. (2011). Fidgetin-Like1 Is a Strong Candidate for a Dynamic Impairment of Male Meiosis Leading to Reduced Testis Weight in Mice. *PLoS One*, in press.

Laissue, P., Burgio, G., l'Hote, D., Renault, G., Marchiol-Fournigault, C., Fradelizi, D., Fellous, M., Serres, C., Montagutelli, X., Monget, P. et al. (2009). Identification of Quantitative Trait Loci responsible for embryonic lethality in mice assessed by ultrasonography. *Int J Dev Biol* **53**, 623-9.

Lander, E. S. and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185-99.

Langer, B., Simeoni, U. and Schlaeder, G. (1998). Prognostic criteria for fetal pyelectasis. *Ultrasound Obstet Gynecol* **11**, 82-3.

Le Du, M. H. and Millan, J. L. (2002). Structural evidence of functional divergence in human alkaline phosphatases. *J Biol Chem* **277**, 49808-14.

Le Du, M. H., Stigbrand, T., Taussig, M. J., Menez, A. and Stura, E. A. (2001). Crystal structure of alkaline phosphatase from human placenta at 1.8 A resolution. Implication for a substrate specificity. *J Biol Chem* **276**, 9158-65.

Leblond, C. P. and Clermont, Y. (1952). Spermiogenesis of rat, mouse, hamster and guinea pig as revealed by the periodic acid-fuchsin sulfurous acid technique. *Am J Anat* **90**, 167-215.

Liu, P. Y., Swerdloff, R. S. and Wang, C. (2010). Recent methodological advances in male hormonal contraception. *Contraception* **82**, 471-5.

Low, N. and Ward, H. (2007). Focus on Chlamydia. *Sex Transm Infect* **83**, 251-2.

Luke-Glaser, S., Pintard, L., Tyers, M. and Peter, M. (2007). The AAA-ATPase FIGL-1 controls mitotic progression, and its levels are regulated by the CUL-3MEL-26 E3 ligase in the C. elegans germ line. *J Cell Sci* **120**, 3179-87.

Mabry, C. C., Bautista, A., Kirk, R. F., Dubilier, L. D., Braunstein, H. and Koepke, J. A. (1970). Familial hyperphosphatase with mental retardation, seizures, and neurologic deficits. *J Pediatr* **77**, 74-85.

Matzuk, M. M. and Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nat Med* **14**, 1197-213.

Misell, L. M., Holochwost, D., Boban, D., Santi, N., Shefi, S., Hellerstein, M. K. and Turek, P. J. (2006). A stable isotope-mass spectrometric method for measuring human spermatogenesis kinetics in vivo. *J Urol* **175**, 242-6; discussion 246.

Moldenhauer, J. S., Ostermeier, G. C., Johnson, A., Diamond, M. P. and Krawetz, S. A. (2003). Diagnosing male factor infertility using microarrays. *J Androl* **24**, 783-9.

Navarro-Costa, P., Plancha, C. E. and Goncalves, J. (2010). Genetic dissection of the AZF regions of the human Y chromosome: thriller or filler for male (in)fertility? *J Biomed Biotechnol* **2010**, 936569.

Nelson, S. M. and Fleming, R. (2007). Obesity and reproduction: impact and interventions. *Curr Opin Obstet Gynecol* **19**, 384-9.

Newman, T. L., Tuzun, E., Morrison, V. A., Hayden, K. E., Ventura, M., McGrath, S. D., Rocchi, M. and Eichler, E. E. (2005). A genome-wide survey of structural variation between human and chimpanzee. *Genome Res* **15**, 1344-56.

Nieschlag, E. (2010). Male hormonal contraception. *Handb Exp Pharmacol*, 197-223.

Nieschlag, E., Vorona, E., Wenk, M., Hemker, A. K., Kamischke, A. and Zitzmann, M. (2011). Hormonal male contraception in men with normal and subnormal semen parameters. *Int J Androl* **34**, 556-67.

Norman, R. J., Noakes, M., Wu, R., Davies, M. J., Moran, L. and Wang, J. X. (2004). Improving reproductive performance in overweight/obese women with effective weight management. *Hum Reprod Update* **10**, 267-80.

Nybo Andersen, A. M., Hansen, K. D., Andersen, P. K. and Davey Smith, G. (2004). Advanced paternal age and risk of fetal death: a cohort study. *Am J Epidemiol* **160**, 1214-22.

Nygren, K. G. and Andersen, A. N. (2001a). Assisted reproductive technology in Europe, 1997. Results generated from European registers by ESHRE. European IVF-Monitoring Programme (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). *Hum Reprod* **16**, 384-91.

Nygren, K. G. and Andersen, A. N. (2001b). Assisted reproductive technology in Europe, 1998. Results generated from European registers by ESHRE. European Society of Human Reproduction and Embryology. *Hum Reprod* **16**, 2459-71.
O'Donnell, L., Nicholls, P. K., O'Bryan, M. K., McLachlan, R. I. and Stanton, P. G. (2011). Spermiation: The process of sperm release. *Spermatogenesis* **1**, 14-35.
O'Donnell, L., Robertson, K. M., Jones, M. E. and Simpson, E. R. (2001). Estrogen and spermatogenesis. *Endocr Rev* **22**, 289-318.
Oku, K., Amengual, O. and Atsumi, T. (2012). Pathophysiology of thrombosis and pregnancy morbidity in the antiphospholipid syndrome. *Eur J Clin Invest*.
Plouffe, L., Jr., White, E. W., Tho, S. P., Sweet, C. S., Layman, L. C., Whitman, G. F. and McDonough, P. G. (1992). Etiologic factors of recurrent abortion and subsequent reproductive performance of couples: have we made any progress in the past 10 years? *Am J Obstet Gynecol* **167**, 313-20; discussion 320-1.
Poongothai, J., Gopenath, T. S. and Manonayaki, S. (2009). Genetics of human male infertility. *Singapore Med J* **50**, 336-47.
Rai, R., Backos, M., Rushworth, F. and Regan, L. (2000). Polycystic ovaries and recurrent miscarriage--a reappraisal. *Hum Reprod* **15**, 612-5.
Rai, R. and Regan, L. (2006). Recurrent miscarriage. *Lancet* **368**, 601-11.
Rai, R. S., Clifford, K., Cohen, H. and Regan, L. (1995). High prospective fetal loss rate in untreated pregnancies of women with recurrent miscarriage and antiphospholipid antibodies. *Hum Reprod* **10**, 3301-4.
Ravel, C., Chantot-Bastaraud, S. and Siffroi, J. P. (2004). [Molecular mechanisms in sex determination: from gene regulation to pathology]. *Gynecol Obstet Fertil* **32**, 584-94.
Regan, L., Braude, P. R. and Trembath, P. L. (1989). Influence of past reproductive performance on risk of spontaneous abortion. *BMJ* **299**, 541-5.
Roussev, R. G., Stern, J. J. and Kaider, B. D. (1998). Anti-endothelial cell antibodies: another cause for pregnancy loss? *Am J Reprod Immunol* **39**, 89-95.
Rull, K., Nagirnaja, L. and Laan, M. (2012). Genetics of recurrent miscarriage: challenges, current knowledge, future directions. *Front Genet* **3**, 34.
Russell, L. B. (1961). Genetics of mammalian sex chromosomes. *Science* **133**, 1795-803.
Salim, R., Regan, L., Woelfer, B., Backos, M. and Jurkovic, D. (2003). A comparative study of the morphology of congenital uterine anomalies in women with and without a history of recurrent first trimester miscarriage. *Hum Reprod* **18**, 162-6.

Schultz, N., Hamra, F. K. and Garbers, D. L. (2003). A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc Natl Acad Sci U S A* **100**, 12201-6.
Sekido, R. and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* **453**, 930-4.
Sharpe, R. M., Kerr, J. B., McKinnell, C. and Millar, M. (1994). Temporal relationship between androgen-dependent changes in the volume of seminiferous tubule fluid, lumen size and seminiferous tubule protein secretion in rats. *J Reprod Fertil* **101**, 193-8.
She, Q. B., Mukherjee, J. J., Chung, T. and Kiss, Z. (2000a). Placental alkaline phosphatase, insulin, and adenine nucleotides or adenosine synergistically promote long-term survival of serum-starved mouse embryo and human fetus fibroblasts. *Cell Signal* **12**, 659-65.
She, Q. B., Mukherjee, J. J., Huang, J. S., Crilly, K. S. and Kiss, Z. (2000b). Growth factor-like effects of placental alkaline phosphatase in human fetus and mouse embryo fibroblasts. *FEBS Lett* **469**, 163-7.
Stephenson, M. D. and Sierra, S. (2006). Reproductive outcomes in recurrent pregnancy loss associated with a parental carrier of a structural chromosome rearrangement. *Hum Reprod* **21**, 1076-82.
Steptoe, P. C. and Edwards, R. G. (1978). Birth after the reimplantation of a human embryo. *Lancet* **2**, 366.
Stevens, V. C. (1997). Some reproductive studies in the baboon. *Hum Reprod Update* **3**, 533-40.
Teklenburg, G., Salker, M., Heijnen, C., Macklon, N. S. and Brosens, J. J. (2010). The molecular basis of recurrent pregnancy loss: impaired natural embryo selection. *Mol Hum Reprod* **16**, 886-95.
Thonneau, P., Marchand, S., Tallec, A., Ferial, M. L., Ducot, B., Lansac, J., Lopes, P., Tabaste, J. M. and Spira, A. (1991). Incidence and main causes of infertility in a resident population (1,850,000) of three French regions (1988-1989). *Hum Reprod* **6**, 811-6.
Tiepolo, L. and Zuffardi, O. (1976). Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Hum Genet* **34**, 119-24.
Tietze, C., Guttmacher, A. F. and Rubin, S. (1950). Time required for conception in 1727 planned pregnancies. *Fertil Steril* **1**, 338-46.
Vatin, M., Burgio, G., Renault, G., Laissue, P., Firlej, V., Mondon, F., Montagutelli, X., Vaiman, D., Serres, C. and Ziyyat, A. (2012). Refined

mapping of a quantitative trait locus on chromosome 1 responsible for mouse embryonic death. *PLoS One* **7**, e43356.

Viudes-de-Castro, M. P. and Vicente, J. S. (1997). Effect of sperm count on the fertility and prolificity rates of meat rabbits. *Anim Reprod Sci* **46**, 313-9.

Waclawska, A. and Kurpisz, M. (2012). Key functional genes of spermatogenesis identified by microarray analysis. *Syst Biol Reprod Med*.

Wennberg, C., Kivela, A. and Holmgren, P. A. (1995). Placental and germ cell alkaline phosphatase RFLPs and haplotypes associated with spontaneous abortion. *Hum Hered* **45**, 272-7.

Wennberg, C., Kozlenkov, A., Di Mauro, S., Frohlander, N., Beckman, L., Hoylaerts, M. F. and Millan, J. L. (2002). Structure, genomic DNA typing, and kinetic characterization of the D allozyme of placental alkaline phosphatase (PLAP/ALPP). *Hum Mutat* **19**, 258-67.

Whyte, M. P., Landt, M., Ryan, L. M., Mulivor, R. A., Henthorn, P. S., Fedde, K. N., Mahuren, J. D. and Coburn, S. P. (1995). Alkaline phosphatase: placental and tissue-nonspecific isoenzymes hydrolyze phosphoethanolamine, inorganic pyrophosphate, and pyridoxal 5'-phosphate. Substrate accumulation in carriers of hypophosphatasia corrects during pregnancy. *J Clin Invest* **95**, 1440-5.

Wilcox, A. J., Weinberg, C. R., O'Connor, J. F., Baird, D. D., Schlatterer, J. P., Canfield, R. E., Armstrong, E. G. and Nisula, B. C. (1988). Incidence of early loss of pregnancy. *N Engl J Med* **319**, 189-94.

Zegers-Hochschild, F., Nygren, K. G., Adamson, G. D., de Mouzon, J., Lancaster, P., Mansour, R. and Sullivan, E. (2006a). The ICMART glossary on ART terminology. *Hum Reprod* **21**, 1968-70.

Zegers-Hochschild, F., Nygren, K. G., Adamson, G. D., de Mouzon, J., Lancaster, P., Mansour, R. and Sullivan, E. (2006b). The International Committee Monitoring Assisted Reproductive Technologies (ICMART) glossary on ART terminology. *Fertil Steril* **86**, 16-9.